

35550

WORLD BANK INSTITUTE

Promoting knowledge and learning for a better world

REFORMING REGIONAL-LOCAL FINANCE IN RUSSIA



Jorge Martinez-Vazquez
Andrey Timofeev
Jameson Boex

Other Titles from the World Bank Institute

India and the Knowledge Economy: Leveraging Strengths and Opportunities

2005. ISBN: 0-8213-6207-0. SKU: 16207

Intergovernmental Finance in Hungary: A Decade of Experience 1990–2000

2005. ISBN: 0-8213-6051-5. SKU: 16051

Intergovernmental Fiscal Relations in Central and Eastern Europe: A Sourcebook and Reference Guide

2005. ISBN: 0-8213-5705-0. SKU: 15705

Reducing Poverty on a Global Scale: Learning and Innovating for Development--Findings from the Shanghai Global Learning Initiative

2005. ISBN: 0-8213-6362-X. SKU: 16362

Beyond Economic Growth: Meeting the Challenges of Global Development (2nd edition)

2004. ISBN: 0-8213-5933-9. SKU: 15933

Building State Capacity in Africa

2004. ISBN: 0-8213-6000-0. SKU: 16000

Granting and Renegotiating Infrastructure Concessions: Doing It Right

2004. ISBN: 0-8213-5792-1. SKU: 15792

Leadership and Innovation in Subnational Government: Case Studies from Latin America

2004. ISBN: 0-8213-5707-7. SKU: 15707

Subnational Data Requirements for Fiscal Decentralization: Case Studies from Central and Eastern Europe

2004. ISBN: 0-8213-5699-2. SKU: 15699

The Right to Tell: The Role of Mass Media in Economic Development

2002. ISBN: 0-8213-5203-2. SKU: 15203

China and the Knowledge Economy: Seizing the 21st Century

2001. ISBN: 0-8213-5005-6. SKU: 15005

Economic Analysis of Investment Operations: Analytical Tools and Practical Applications

2001. ISBN: 0-8213-4850-7. SKU: 14850

WBI LEARNING RESOURCES SERIES

Reforming Regional-Local Finance in Russia

Jorge Martinez-Vazquez, Andrey Timofeev, and Jameson Boex

The World Bank
Washington, D.C.

©2006 The International Bank for Reconstruction and Development / The World Bank
1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org
E-mail: feedback@worldbank.org

All rights reserved

1 2 3 4 5 09 08 07 06

This volume is a product of the staff of the International Bank for Reconstruction and Development / The World Bank. The findings, interpretations, and conclusions expressed in this volume do not necessarily reflect the views of the Executive Directors of The World Bank or the governments they represent.

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

Rights and Permissions

The material in this publication is copyrighted. Copying and/or transmitting portions or all of this work without permission may be a violation of applicable law. The International Bank for Reconstruction and Development / The World Bank encourages dissemination of its work and will normally grant permission to reproduce portions of the work promptly.

For permission to photocopy or reprint any part of this work, please send a request with complete information to the Copyright Clearance Center Inc., 222 Rosewood Drive, Danvers, MA 01923, USA; telephone: 978-750-8400; fax: 978-750-4470; Internet: www.copyright.com.

All other queries on rights and licenses, including subsidiary rights, should be addressed to the Office of the Publisher, The World Bank, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2422; e-mail: pubrights@worldbank.org.

ISBN-10: 0-8213-6557-6
ISBN-13: 978-0-8213-6557-1
eISBN: 0-8213-6558-4 (e-book)
DOI: 10.1596/978-0-8213-6557-1

Library of Congress Cataloging-in-Publication Data

Martinez-Vazquez, Jorge.

Reforming regional-local finance in Russia / Jorge Martinez-Vazquez, Andrey

Timofeev, Jameson Boex.

p. cm. — (WBI learning resources series)

“This book complements the earlier WBI title *Russia’s transition to a new federalism*” —Foreword.

Includes bibliographical references.

Contents: Contents: Introduction — A historical overview of Russia’s transition — The territorial and administrative structure of local governments — Subnational government responsibilities — The assignment of revenues at the subnational level — Intergovernmental fiscal transfers — Subnational borrowing — Budgeting practices and fiscal management at the subnational level — Conclusion.

ISBN-13: 978-0-8213-6557-1

ISBN-10: 0-8213-6557-6

ISBN-13: 978-0-8213-6558-8 (e-book)

ISBN-10: 0-8213-6558-4 (e-book)

1. Intergovernmental fiscal relations—Russia (Federation) 2. Local finance—Russia (Federation) 3. Finance, Public—Russia (Federation) I. Timofeev, Andrey. II. Boex, Jameson, 1971-III. Title. IV. Series.

HJ1211.52.Z7M36 2006
352.4'2140947--dc22

2005057898

Contents

Foreword v

1. Introduction 1
 2. A Historical Overview of Russia's Transition 5
 3. The Territorial and Administrative Structure of Local Governments 15
 4. Subnational Government Responsibilities 57
 5. The Assignment of Revenues at the Subnational Level 81
 6. Intergovernmental Fiscal Transfers 145
 7. Subnational Borrowing 171
 8. Budgeting Practices and Fiscal Management at the Subnational Level 185
 9. Conclusion 201
- References 207

Foreword

Reforming Regional and Local Finance in Russia assesses the current state of Russia's regional and local fiscal relations and self-government, and analyzes the related reforms undertaken over the past decade. The analysis is based on the kind of first-hand experience and in-depth knowledge of Russian intergovernmental reforms that only a handful of Western scholars possess. This book complements the earlier WBI title *Russia's Transition to a New Federalism* (2001), reflecting the shift in Russia's policy reform agenda from an emphasis on federal and regional relations to the regional and local levels.

The study uses a framework that integrates all the building blocks of federalism: size and structure of jurisdictions, expenditures, revenues, transfers, and borrowing. It offers reform options based on international practices and normative principles, while also identifying some dangers that may arise in implementing the next round of proposed intergovernmental reforms in Russia. The book will be of particular interest to policymakers, researchers, and teachers.

Reforming Regional and Local Finance in Russia was prepared as part of the World Bank Institute's program to train central and local government officials to carry out intergovernmental reforms, and to build a core group of local trainers who can deliver future programs independently. I am pleased to make this volume available to all who are engaged or interested in intergovernmental reforms in Russia as well as in Central Asia and the Caucasus.

Frannie A. Léautier
Vice President
The World Bank Institute

1

Introduction

Although the regional-local dimension of public finance and intergovernmental relations is of high concern for the Russian federal government, it is often overlooked both by federal policy-makers and analysts and observers of fiscal federalism in the Russian Federation. For example, while authority over tax policy is practically monopolized by the federal government in Russia, the federal *Tax Code* of 1999 actually fails to accommodate local tax collections under the different local government structures established within the Russian regions. In fact, the federal Ministry of Finance does not directly deal with local governments and has to rely on information received from regional governments. While the regional level mostly provides liaison between the federal and local authorities, local governments in fact carry out the bulk of social service provision (education, health, and social welfare). As a result, most of the subnational fiscal problems that the federal government has to solve (largely by means of bailouts of housing and utilities) are concentrated at the local level. As the federal government is constitutionally constrained in addressing the local sector directly, it has to rely on the development of a general framework for regional-local relations that would embed points of entry for federal policies and provide incentives to regional and local governments to enact desirable institutional and policy adjustments.

As noted in *Russia's Transition to a New Federalism* (Martinez-Vazquez and Boex 2001), considerable progress has been made since the start of Russia's transition in the rationalization of intergovernmental relations at the federal-regional levels. However, intergovernmental fiscal relations at the regional-local levels have seen much less progress. Although there have been some reform efforts at the regional level, many Russian regional and local governments continue to employ budgetary practices inherited from the Soviet system. In many regions, public servants continue to think and act as they did under the previous regime, for example, by not taking initiative and passively waiting for guidelines from a higher authority. The roles of public officials at the regional and local levels are only slowly being modified to adapt to a more decentralized institutional environment. For example, one of the main roles of finance officers under the previous system was to aggregate budget requests from constituent governments and public entities and pass them on for funding to the higher authority. However, when the federal authorities stopped considering actual budgetary expenditures and switched to more appropriate indicators of expenditure needs in financing lower-level governments, regional governments had to revise their relations with localities.

The current approach of the federal government toward regional-local relations presents a mix of rigid norms and loose guidelines. The Constitution gives the federal and regional governments joint responsibility for the implementation of the citizens' right to local self-government. This has enabled the federal government to pass a number of constitutional laws that develop the framework for local governance. Regional governments were supposed to establish the system of local self-government on their territory in accordance with federal legislation by passing regional legislation that would further elaborate the regional-local government arrangements. However, in areas where regional governments had some degree of policy freedom, often many did not have the required technical expertise to make their own informed policy choices. The easiest solution for defining regional-local relations seems to have been to copy the federal approach to federal-regional relations and apply it at the regional-local level.

However, in a number of key aspects the regional-local sector differs from the federal-regional sector, while regional-local relations also exhibit significant variation across regions. For example, on average, the current decentralization of expenditures is higher in the regional-local sector than in the federal-regional sector. At the same time, the share of own and assigned revenue is larger at the regional level than in local budgets. Dependence of many regions on federal grants further aggravates vertical imbalances within regions. Moreover, unlike constitutionally equal subjects of the Federation, localities differ by type: urban versus rural and territorial-based versus settlement-based. Another (technical but important) issue in regional-local relations is that, unlike at the federal-regional level, objective data are not available for many important local characteristics that would be good candidates for regional equalization formulas. These data restrictions also limit the extent of quantitative analysis in this book.

Given the great diversity of regional conditions in the Russian Federation, it seems appropriate that the federal policy toward regional-local relations provides a balance between flexibility and uniformity. Thus, in addition to the familiar mechanism of inducing desired behaviors through federal transfers, federal guidelines could be introduced or recommended to the regional governments containing several suitable blueprints with alternative frameworks for regional-local relations. Ideally, at least some of these alternatives would come from successful experiments by the Russian regions themselves. However, before such experience is accumulated, normative guidance and experience and lessons from intergovernmental relations at the subnational level in other countries can be of some value.

The main goal of this book is to assess the current state of regional-local fiscal relations and local self-government in Russia. This assessment enables us to suggest options for reforms based on international practices and normative principles, and gives us cause for concern about some dangers contained in the pending reform proposals.

By focusing mainly on subnational finance in the Russian Federation, this book is intended to complement and in some ways update *Russia's Transition to a New Federalism* (Martinez-Vazquez and Boex 2001). That earlier study focused on federal-regional relations, which were of crucial importance for Russia's survival as a country and its fiscal balance throughout the 1990s. The new millennium seems to have brought resolution to many of those problems, many of which were of a political nature. Although some of these hard issues have not yet been fully resolved, the policy focus has shifted to the issues of public service delivery and technical efficiency in the public sector, where most of the problem areas ended up at the local level as a result of the "offloading" of expenditure responsibilities in the early 1990s. Moreover, the success of local government reforms has bearings for the future of Russia as a democratic state. As political scientists frequently argue, local government is an ideal "school of democracy"; as citizens are made more responsible for local affairs, the opportunities for participation are enhanced and feelings of empowerment are fostered (see, for example, Sharpe 1970). In addition, the achievement of viable local self-governance could anchor a new balance of power among levels of government in Russia and perhaps bring some balance to the current trend for recentralization in federal-regional relations.

The broad scope of the topic under consideration requires a systematic organizational structure, sequentially addressing the key dimensions of regional-local relations. Therefore whenever detailed policy issues arise on specific topics, they are discussed in appendixes to the individual chapters. The rest of the book is organized as follows:

In the next chapter (chapter 2) we review the political background in which local self-government evolved in Russia in the course of the previous decade. We show how political developments shaped the present fiscal arrangements and also discuss the new trends that set the stage for further reforms.

The subject of chapter 3 is the administrative division of local jurisdictions and the internal organization of local governments, including civil service. Organizational structure underpins

the institutional set-up under which fiscal decentralization will be implemented and administered.

In chapter 4 we assess the scope of responsibilities devolved to the local level and the variation in this devolution both across and within regions. Getting the assignment of responsibilities right at the local, regional, and central levels is fundamental for improving efficiency and fairness and for increasing the accountability of government to the citizens.

In chapter 5 we examine how the devolved responsibilities are matched with revenue-raising powers authorized for local governments and the resulting degree of revenue autonomy. Revenue assignments should be compatible with the rest of the decentralization system in terms of vertical balance and with the ongoing and planned reforms in expenditure assignments, transfers, and budget process.

Chapter 6 addresses vertical and horizontal fiscal imbalances at the local level and the design and implementation of intergovernmental transfers at the regional-local level. Properly designed transfer systems and revenue assignments can help compensate for financial insufficiencies and economic disparities across jurisdictions and for imbalances among different levels of governments.

In chapter 7 we turn to the issue of subnational borrowing. The evolution of the Russian legal framework for subnational government borrowing shows a gradual introduction of external constraints on regional and local governments' borrowing powers. One of the questions we seek to answer in that chapter is whether the new policy provides the right balance between fiscal autonomy and responsiveness.

The focus of chapter 8 is on budgeting practices of subnational governments in Russia. We take a look at best practices and principles in fiscal management around the world to identify policies and reforms that could be introduced to improve budgeting at the subnational level in Russia.

In chapter 9 we summarize our findings, evaluate the pending reform proposals, and point out the most important issues in regional-local relations that Russia still has to address.

2

A Historical Overview of Russia's Transition

The current Russian system of government is heavily based on the hierarchical system of administrative subordination inherited from the Soviet Union.¹ The government structure of Soviet Russia had four tiers of government: (1) the Russian Soviet Federative Socialist Republic (RSFSR); (2) the regional tier (ethnic republics, *krais*, *okrugs*, *oblasts*, and autonomous areas);² (3) a first local tier (subregion cities and *rayons*);³ (4) a second local tier (subcity districts and subrayon towns, townships, and rural districts). Before the *perestroika* era, each administrative unit was governed by two parallel bodies: the local committee of the Communist Party and the local Soviet (council). However, elections to the Soviets were not competitive and a single candidate for each district was effectively nominated by the Party. Thus, all decisionmaking was handled within the Party apparatus and local Soviets only legitimized these decisions and implemented them through the local executive branch.

The transformation of subnational governance started in June 1987, when several rayons were allowed to hold multicandidate elections to local Soviets. The next subnational elections of March 1990 provided for competition among several candidates for each seat in all subnational Soviets. This introduced some horizontal accountability (in other words, accountability of local officials to their constituencies) in the subnational tiers of government in an otherwise still highly centralized administrative hierarchy. In fact, *perestroika* attempted to revive the old Bolshevik slogan: "All powers to the Soviets!" Thus, popularly elected Deputies of Soviets began appointing the executive branch at all levels of government. Russia maintained this system of government after the dissolution of the Soviet Union until the adoption of the new Constitution in 1993.

Political Transition and Just Coping from 1989 to 1998

The Russian Federation was born in the agony of the disintegration of the Soviet Union, thereby marking the young country from the start with the fear of its own disintegration. These fears were not without justification: the Russian Federation was formed of 89 very different regions. From the start, several of these regions rushed to declare their sovereignty and independence from the Russian Federation. In many ways, therefore, the new country was subject to centrifugal forces similar to those that had led to the disintegration of the Soviet Union.

In the early years of the Boris Yeltsin presidency, through the mid-1990s, intergovernmental fiscal relations were dominated by pressures to contain and mitigate powerful centrifugal forces in the country. Many regions were attempting to position themselves to benefit as much as possible from the political and institutional weaknesses of the center during this early period of the transition. These same regions often flaunted federal laws and by so doing, imposed explicit costs and negative externalities on other regions. The desire of the federal authorities to

1. This section draws from Martinez-Vazquez (forthcoming).

2. *Oblast* is a name for the subnational entity in several Slavic languages. The word *krai* (which also means *border* or *end*), is used for regions located along the economic and geographic periphery. *Okrug* is a Slavic loan translation of German *Kreis*, a term to denote administrative subdivision.

3. A *rayon* is the Russian equivalent of a local government district, such as a U.S. county.

find conditions acceptable to troublemaking regions led inevitably to different forms of fiscal policy—or asymmetric fiscal federalism.⁴

The most intense demands for special (political and fiscal) treatment were coming from ethnic republics with separatist tendencies as expressed by the declarations of independence in Tatarstan, Bashkiria, and Chechnya. While in the case of the former two, the situation was managed through the negotiation of treaties with the federal government as early as 1993, in the case of Chechnya it led to a bloody civil war. A number of natural resource-rich regions, for example Sakha (Yakutia) and Tyumen, demanded special arrangements and treatment on the basis that these regions had been exploited for their natural resources, had been subjected to environmental degradation, and had otherwise never benefited from their mineral wealth. Less aggressive in their demands, at least at the beginning of the transition, were the industrially well-endowed and in general wealthier regions including Moscow City and Sverdlovsk Oblast. During the second half of the 1990s, these regions would join the chorus of complaining regions, in their case protesting that they were being forced to subsidize many other, poorer regions.

The federal government in the early years of the transition remained reactive and tried to adapt to the agenda set by the maverick regions. It should be noted that the federal government during 1992–93 did not have much of a choice but to accept the reality of an asymmetric system of intergovernmental relations, which was being fast shaped by the demands of a small number of ethnic republics. In fact, it was the regions that early on dictated the agenda of the federal contract—the shape and form the Russian Federation should assume. In stark contrast to the central control during the Soviet era, during the early transition period regions were able to take advantage of the fact that they had *de facto* control over key elements of government administration, including tax administration and internal security. This left the federal government without tools to deal with noncompliant regions and enforce federal legislation.

Although the ethnic republics were given some additional powers in the Federation Treaty of March 1992, these powers were taken away by the 1993 Constitution. Regional assertiveness, especially on the side of the ethnic republics, was at peak during 1992–93 as the President and the Supreme Soviet (Parliament) were involved in a power struggle. The standoff between the President and the Supreme Soviet culminated in the shelling of the Supreme Soviet building and the dissolution of the elected government bodies at all levels of government throughout the country. Until the adoption of the new Constitution with the 1993 referendum, the nation was governed by (a hierarchy of) appointees.

The Russian Constitution of 1993 recognized the possibility of an asymmetric configuration of intergovernmental relations between the regions and the federal government by virtue of allowing bilateral (federal-regional) treaties. However, following its usual principle of checks and balances, Yeltsin's administration also lobbied for constitutional provisions for local self-government, thus trying to challenge the power of the then-mighty regional governors. According to the new Constitution, all power is vested in the people and exercised by them directly or indirectly through a system of representation involving regular free elections. The Constitution guarantees local communities the right to govern local affairs separately from the sovereignty

4. It must be noted that asymmetric federalism was to some extent practiced in the Soviet Union since subnational budgets were customized on the basis of negotiated expenditure norms, "regulated" revenue sharing rates, and bargained transfers. The most visible form of asymmetric treatment was the regulation of taxes, which had the objective of reducing funding to richer regions. However, during the early transition, regulation was seen by many as an opportunity for the rich and powerful regions to benefit through their large bargaining power. In retrospect, at least in the early years of the transition, regulation of taxes led to a high degree of equal expenditures per capita across regions (see Martinez-Vazquez and Boex 2001). But besides being an effective equalization tool, tax regulation presented many problems, among them creating perverse incentives for revenue mobilization by subnational governments.

of the state.⁵ The 1995 *Law on the General Principles of Organization of Local Self-Government in the Russian Federation* further developed the concept of autonomy of local self-government.⁶ Later, the Presidential Administration also sponsored two associations of local governments: the Congress of Municipal Settlements⁷ and the Union of Russian Cities.

While providing a structural framework for federal-subnational relations, the formal acceptance of asymmetric relations through bilateral treaties in the 1993 Constitution brought back some predictability and order to the system of intergovernmental fiscal relations.⁸ While confrontations between the federal government and some regions continued to test the strength of the federation during this period, the nature of these confrontations tended to shift from centrifugal tensions toward a competition between regions for special recognition and favorable fiscal arrangements. While the bilateral treaties provided an official acknowledgment of regional power, the general pattern of behavior among the regions became less chaotic and threatening to Russia's unity (more on this can be found in Martinez-Vazquez forthcoming). Of course, the very notable exception to all this was Chechnya. Another sign of regional power during this period, and one that is particularly important for this study, was that regional governments had absolute discretion to organize their relationships with their local governments.

In the eyes of many, the asymmetric design of the system of intergovernmental fiscal relations saved Russia from falling into an abyss of civil wars as in Chechnya and thus kept the country from disintegrating. But asymmetric federalism has not been without costs in the Russian Federation. Besides generating mistrust and resentment among the regions, the practice of asymmetric federalism in general and the bilateral treaties in particular had other significant costs. One of them was the mounting fiscal pressure on the federal budget, which became increasingly noticeable during the first six months of 1998.⁹ The conflicting budgetary demands at the regional and federal levels and the inability of the federal governments to collect revenues eventually meant a federal deficit as high as 10 percent of gross domestic product (GDP). At the same time, the federal and regional governments both continued to borrow more heavily domestically and abroad. All levels of government proceeded to accumulate payment arrears while they also became more frequent users of non-cash offsets and barter. The combination of increased deficits with a tight money supply and fixed exchange rates led to the August 1998 crisis, with the devaluation and floating of the ruble and the default by the federal government in most of its domestic debt, which in turn precipitated a banking crisis.

Many observers have concluded that the problems with Russian federalism during these times were a manifestation of a weak federal government, the absence of cooperation between the center and the regions, and a "common pool" problem. This is a main theme in Blanchard

5. Articles 130 through 133 of the Constitution grant local communities autonomy in governing local affairs, planning and executing local budgets, introducing local taxes and surcharges, managing municipal property, securing public order, and other spheres. This right is to be exercised through referenda, elections, and other forms of direct vote, as well as councils and other bodies of self-government. Changes to local authorities' boundaries are allowed only after consultations with the communities concerned.

6. Article 6 determines the scope of local affairs; gives the inhabitants of any settlement, urban or rural, the right to local self-government regardless of the settlement's size; prohibits subordination of one municipality to another; and requires that a regional law separate powers of two municipalities when one is located inside the other. Article 36 prohibits regional governments to claw back fiscal surplus resulting in local budgets. Article 42 grants local governments virtually unconstrained access to capital markets including establishment of municipal institutions of credit and finance. Articles 47–48 establish the accountability of local governments to their constituency via confidence votes and to individuals and legal entities in courts.

7. Presidential Decree No. 1281 of October 1998 on the Congress of Municipal Settlements.

8. Bilateral agreements covered issues on budgetary relations, state property, ownership and use of natural resources, regional migration, and so on.

9. See, for example, World Bank (2001).

and Shleifer (2000); Lavrov, Litwack, and Sutherland (2000); and OECD (2000). Blanchard and Shleifer, in particular, emphasized that the contrast between the “success” of China’s transition and the “failure” of Russia’s lay in the fact that China was able to retain a strong political center. Mending Russia’s fiscal federalism problems required strengthening the political and institutional power of the center relative to the regions.¹⁰ Russia would have to wait until the election of Vladimir Putin as president for a significant turnaround in the relationships between the center and the regions. But the situation may have involved more than a weak center. For example, de Figueiredo and Weingast (2001) have added that the problem with Russian federalism was also the lack of “appropriately defined limits on the central government.” The federal government in Russia was in some ways “too strong,” given its ability to change rules and extract rents from the regions, which detracted from the benefits regions derived from participation in the federal relationship. From this perspective, solving the problem of noncooperation in Russia required not only more power for the center but simultaneously credible limits on the center, which had acted and could continue to act without self-restraint in the pursuit of its own interests.¹¹

Consolidation of Power since 1998

Many of the problems and tribulations Russia went through in the last decade—lack of fiscal discipline, economic stagnation, and so on—have been explained by many observers as having roots in the inability of the federal government to impose a unified legal system throughout the Russian Federation. After almost a decade of transition, Russia’s federalism is being slowly but surely returned to a symmetric base, a trend that started in the closing years of the Yeltsin presidency and accelerated after Putin became president. The administration of President Yeltsin clearly started a campaign during 1997–98 to rein in the regions. This policy was significantly intensified by the new administration of President Putin, who, since he took over in 2000, has made gaining control over the regions and enforcing the federal laws a cornerstone of his administration. In fact, one of Putin’s first prominent moves was to cut back the power of regional governors, including assigning himself authority to remove incompliant governors from office.

The crisis of August 1998 provided an opportunity for the federal government to reevaluate its policies on many fronts, including its budgetary relations with the regions. One of the issues examined was the role played in the crisis by the bilateral treaties. It became clear that the special fiscal treatment provisions in the treaties had contributed to the mounting fiscal pressures that eventually led to the crisis. Undoing the damage from bilateral treaties and the special deals proved to be difficult. Nevertheless, the retreat from asymmetric treatment clearly started in the late Yeltsin years by simply not always complying with the provisions in the treaties. At that point the federal government also started the serious recentralization of fiscal resources.¹² These efforts involve the introduction of deeper tax reforms, including eventually a *Tax Code*, a *Budget Code*, the *Law on the Financial Foundations of Local Self-Government in the Russian Federation*, and the “Concept of Reform of Inter-Governmental Relations in the Russian Federation for 1999–2001.” For example, another law, the *Law on the Principles for the Demarcation of Jurisdictions and Powers*, enacted in June 1999, clearly restated the supremacy of the federal constitution, federal legislation, and federal decrees over regional constitutions, legislation, and

10. See also the discussion in Martinez-Vazquez and Boex (2001) and Shleifer and Treisman (2000).

11. For example, at the start of the transition the federal government dumped expenditure responsibilities on the regions without adequate financing. The federal government has continued to control basically all tax and regulatory powers. Since the regions depend on the center for most of their budget funding, they continue to be subject to the whims of federal authorities.

12. See Martinez-Vazquez and Boex (2001).

decrees. One clear feature of all of these documents, and one that is clearly important for this study, was the new willingness of the federal government to intervene in the fiscal arrangements between regional governments and their local governments.

The 1997 *Law on the Financial Foundations of Local Self-Government in the Russian Federation* brought about some structure to intraregional fiscal relations, over which regional governments hitherto had had unlimited discretion.¹³ The law also put constraints¹⁴ on local finance in an effort to rein in the borrowing power of local governments. More recently, the *Budget and Tax Codes* included several other important restrictions.¹⁵ Moreover, there have been signs in recent times that the federal government is trying to curb the political autonomy of local governments. Thus, the amendment of August 2000 to the *Law on the General Principles of Organization of Local Self-Government in the Russian Federation* established the accountability of local officials to upper-level governments. The amendment allows a regional governor (or the President of the Russian Federation) to dissolve a local council or dismiss a municipal head for failure to recall an act ruled by court to contradict federal or regional laws or for abusing human or citizen rights. One interpretation of this new political stance is that the new administration does not need local self-governance any longer to balance the powers of regional authorities. An associated interpretation is that the federal center might wish to extend its grip on power to the local level after the regional authorities have been essentially tamed. Indeed, about the same time the State Duma approved legislation that gave the Russian president power to suspend regional legislation that conflicted with federal law, as well as power to dismiss regional governors if their actions were judged to be in violation of federal statutes. Some have seen these moves as the federal center's attempts to restore the "Russian matrushka-doll" type of hierarchy of power.

President Putin's election in March 2000—following Yeltsin's retirement in December 1999—gave him what now seems a clear mandate to control the regional governments and reassert the role of the federal authorities. Putin was elected with a wide margin and his popularity allowed him to immediately take on the regional governors and reduce their power. For example, the chief prosecutor's office declared in May 2000 that there were 60 regions with local laws seriously contradicting federal laws, including the regional charters or constitutions. In June 2000 the chief prosecutor (appointed by Putin) gave the regions one month to synchronize their laws with the federal laws. And this was just the beginning.

In a further effort to assert federal control, Putin issued a decree in May 2000 that divided Russia into seven groups of regions, each with a presidential envoy to monitor regional legislation and ensure that regional administrations were abiding by federal laws. These envoys replaced 80 presidential regional representatives who had been appointed by a 1997 presidential decree and had largely turned into ineffective figureheads. Political reactions to the decree varied from fear by regional governments concerned with losing regional powers, to skepticism and concern that the decree would merely impose an additional layer of ineffective federal

13. Ambiguity of legislation on local self-government brought at times significant differences in implementation across regions. Thus, while a majority of regions have local jurisdictions based on two administrative tiers (cities and rayons) of the Soviet government structure, several oblasts in the Urals established self-government in every settlement.

14. Article 13 of the law bans the transfer of municipal property as a share of the charter capital of credit institutions. Article 16 of the law states that bonds may be issued only for investment purposes, restricts municipal debt to 15 percent of budgetary expenditures, and specifies that local borrowing is not guaranteed by the State.

15. Regional and local governments may collect revenues only from federally assigned taxes, must fit their bases to the federal law, and may levy rates only within federal limits. The *Budget Code* limits the deficit of local budgets to 3 percent of pretransfer revenues. The overall level of municipal debt is constrained in the *Budget Code* by capping expenditures for debt service in any year at 15 percent of budgetary expenditures.

bureaucracy. It appears that the new administrative controls, in concert with the legislative measures discussed below, contributed to increased regional compliance with federal laws.

A further legislative reform enacted in 2000 restructured the Federation Council (the upper chamber of Parliament) by replacing the governors with regional representatives appointed by the regional legislatures. In the past, the regional governors as members of the Federation Council had often played an obstructionist role. This situation was exacerbated by the fact that regional governors had parliamentary immunity, as their position automatically gave them a seat in the Federation Council, Russia's upper house of Parliament.

The legal changes gave the federal authorities effective instruments even to deal with the most recalcitrant regions, excluding, of course, war-torn Chechnya. The best evidence that the federal government has regained much authority is that even Tatarstan moved to adopt amendments to its Constitution to bring it closer in line with the Russian Constitution and other federal legislation.¹⁶ But not all is clear; Bashkortostan, even under pressure from Moscow, is holding on to its declaration of sovereignty and its President Murtaza Rakhimov has run a third time for office on the "sovereignty" platform.¹⁷

Putin's administration continued the recentralization of revenues that had started in the late years of the Yeltsin administration. The initial effort was the reform of the Russian tax system in 1998 as part of the emergency measures taken in response to the economic crisis of August 1998. The first part of the new *Tax Code*, which contains the general principles of taxation in the Russian Federation, was passed by the legislature in 1999. Subsequently, however, the government's tax reform efforts effectively stalled as the more important, second part of the *Tax Code*, which contained legislation covering individual taxes, became bogged down in the legislative process because of a lack of political consensus on the nature of the reforms within the executive and legislative branches. Following his election to the presidency, Putin managed to get four chapters of the second part of the *Tax Code* approved by the State Duma, which had a profound impact on the regional finances. Subnational government turnover taxes, which provided a substantial level of own-source revenues for local and regional governments, were eventually eliminated.

The cumulative impact of the post-1998 tax reforms and the reassignment of revenue sources on the vertical fiscal balance is reflected by the changing distribution of resources across the levels of government over time. During the early years of the transition the federal share of overall tax revenues steadily declined. While in 1992 approximately 60 percent of overall tax collections were assigned to the federal level, by 1997 the federal share had declined to 42.5 percent. A sharp reversal in this trend began in 1998; since its low point in 1997, the federal share of consolidated tax revenues has steadily increased over time. The federal government's share of overall tax collections once again reached 60 percent by 2001 and further increased to 65.1 percent in 2002.¹⁸ Thus, after a period of decentralization of fiscal resources early in the

16. See the "Russian Regional Report" of the East-West Institute (Vol. 7, No. 7, Feb 20, 2002).

17. See the "Russian Regional Report" of the East-West Institute (Vol. 7, No. 12, March 27, 2002).

18. One can spot the similarity to the developments in China before and after the 1994 reform. Both countries had practiced *de facto* upward collection of taxes and negotiated sharing of revenue. Both eventually centralized VAT revenue and devolved personal income tax (PIT) revenue while sharing the corporate income tax (CIT) base (by enterprise ownership in China and through supplemental rates in Russia). The crucial difference is that in China subnational governments were granted tax administration authorities over taxes whose revenue they are exclusively entitled to while in Russia the centralized tax administration of all taxes was strengthened. There is also a remarkable difference in institutional development: China continues to be a single-party regime and still has to undergo deep tax and budget reforms while Russia is a working democracy which has introduced in recent years advanced *Tax and Budget Codes* (for more on China see Bahl 1999; Bahl and Martinez-Vazquez forthcoming; and Martinez-Vazquez and Wong 2003).

transition, in recent years the revenue balance in Russia has shifted substantially in the federal government's favor.¹⁹

Overall, the tax reforms have tended to reduce the complexity of the tax system and lower the marginal tax rates, thereby reducing incentives for tax avoidance and evasion. While generally in line with sound fiscal practices, these reforms and changes in the assignments of tax revenues have caused a substantial loss of shared revenues for regional and local governments and have added very little to the revenue autonomy of these governments. In contrast, federal revenue sources, such as excise taxes, have been substantially increased as part of ongoing tax reforms. In this sense, the tax reforms did not facilitate, and in fact may jeopardize, many of the potential benefits of a decentralized fiscal system.

Some improvements have been achieved on the expenditure side as well. Article 85 of the *Budget Code* was modified in 2000 to require that joint expenditure responsibilities must be listed separately in the annual budget law for every type of activity. For those federal mandates that the federal government is unable or unwilling to fund, changes to Article 83 of the *Budget Code* (which took effect for the 2001 budget year) require that the annual budget law include the list of mandates that will be stricken because of a lack of funding. For mandates that are partially funded, the budget law will need to identify specifically which parts of the mandates would not be financed by the federal government. A number of further changes to the *Budget Code* aimed to improve the budget process by clarifying sanctions for violating federal budget laws and imposing additional limits and regulations for subnational government borrowing. However, despite these improvements, the current reforms ultimately fail to address the most important cause of vertical fiscal imbalance: the lack of tax autonomy for regional and local governments.

What Does the Future Hold?

Russia has made a remarkable journey over the past decade. It has returned from the brink of disintegration and chaos to become a promising young democracy. The challenge for the future remains, however—to find the appropriate balance between a federal government capable of enforcing the law and protecting and defending common national interests and the rights of the regions, including respect for and acceptance of their diversity. Moreover, the new system of governance should provide some space for independent local self-government. The latter needs to include a national consensus on what limits federal legislation should impose on the political and fiscal relationships between the regions and their local governments.

An important recent development in fiscal federalism has been the appointment by President Putin of a high-level commission in 2001, which became known as the Kozak Commission, for redefining and reforming intergovernmental fiscal relations in Russia.²⁰ The Commission received substantial input from fiscal experts and think tanks in Russia and engaged in considerable discussion before reaching its final conclusions. The Commission issued recommendations in September 2002,²¹ but President Putin decided to send the recommendations to the State Council for further consideration before sending any bills to the Parliament (State

19. The determination of the "right" level of revenue centralization is a policy decision that balances competing policy objectives, including macroeconomic stability and administrative efficiency, which often require centralized control over fiscal policy tools, and the desire to provide subnational governments with own-source revenues to fund the decentralized provision of public goods and increased efficiency and accountability of subnational governments.

20. Dmitrii Kozak is deputy head of the President's Administration.

21. Originally the Commission was supposed to have submitted its findings and proposals by early summer 2002. The deadline was subsequently pushed back.

Duma).²² Even after the State Council formally approved the gist of the reform in October 2002, the Duma vote had to be postponed from December 2002 to February 2003 to garner sufficient support.

The recommendations of the Kozak Commission appear to cover the entire spectrum of intergovernmental relations. Among other things the recommendations include the need to fund each level of government adequately according to the expenditure responsibilities they have been assigned. This may be interpreted as more money for subnational governments—in particular, retention of a larger share of tax revenues at the point of collection. The increase in tax-revenue retention will not negate the need for transfers, however, since only a tiny minority of Russia's 13,000 local governments can be considered "surplus jurisdictions" that have enough money to meet or exceed their expenses with current revenue assignments.²³

One of the most divisive issues is the reassignment of revenues for taxes from natural resources. This was a hotly debated issue in 2002, when taxation of mineral resources was changed from the subsoil use fees to the extraction tax, so that revenue retention was eliminated for local governments and reduced for regional governments. For the hydrocarbon extractions, the new revenue split between the federal and regional levels was set at 80:20 instead of 60:40. The Kozak Commission report appears to argue that these revenues should be recentralized completely and then redistributed more fairly among the regions.²⁴ The amendments to the *Tax and Budget Codes* proposed by the Ministry of Finance allocate the revenue from the natural resources extraction tax between the federal and regional governments in proportions 100:0 for gas and 90:10 for oil. Furthermore, the Kozak Commission recommended abandoning the "two key" practice, which has required that federal and regional governments must issue exploitation licenses for natural resources, for full and exclusive federal control. There is no need to describe the kind of opposition these proposals have brought from the natural resource-rich regions.

The presidential administration introduced two laws, respectively, on the reform of regional and local government to the State Duma in early January 2003 and simultaneously sent them for discussion in the regions. With the intention of backing these bills, President Putin argued that ". . . Russia no longer had time for demagoguery and empty promises."²⁵ The fact that Putin associated himself with local government reform with about one year remaining before the presidential elections suggested at that time that he would work to ensure that this legislation would go through both houses of the federal legislature, despite the expected strong resistance from the regional governors.

Indeed, in February 2003, the bills were approved by the lower house of Parliament in the first (or framework) reading. However, the deputies requested the government to present associated changes in the tax assignment before they would consider the bills in the second (essential) reading. Putin ordered the government to prepare amendments to the *Tax and Budget Codes* introducing necessary reassignment of revenue sources by April 1, 2003. However, the government missed that deadline and the President moved it ahead to early June. Despite the extended deadline, the government once again failed to finalize the new revenue assignment,

22. The State Council is a consultative body whose membership includes Putin and rotated representation of Russia's 89 governors. This body partially serves as a sweetener for governors who had to leave the federal scene after the reform of the upper chamber of Parliament, the Federation Council.

23. This is according to Oleg Sysuev, president of the Congress of Local Governments as reported in the East-West Institute's "Russian Regional Report" (Vol. 7, No. 28, September 27, 2002). When tax-sharing rates were "regulated" in the early years of the transition it was not uncommon to let local governments keep 100 percent of all taxes collected in the jurisdiction, including such main taxes as the VAT and income taxes. Even then it was necessary to implement subsidies or transfers.

24. Currently six regions receive close to half of all revenues from the exploitation of natural resources.

25. See the "Russian Regional Report" of the East-West Institute (Vol. 8, No. 4, March 14, 2003).

arguing that it had to know the Parliament's final delineation of responsibilities in the law. Instead, on June 1, the government sent to the Parliament a preliminary draft of amendments to the *Tax* and *Budget Codes*. Nevertheless, with pressure from the presidential administration, in September 2003 the deputies passed both the laws in the third (final) reading, without the associated revenue assignment.

It is uncertain where President Putin and his administration actually want to take fiscal federalism, but it seems that more central control (from Moscow) and a strengthened federal structure are high on the policy agenda. This does not necessarily mean that there will be sustained fiscal recentralization in Russia. Perhaps the current administration believes that in order to achieve effective decentralization it is necessary to forge a stronger federal government, one that is able to enforce the law throughout the national territory. However, Putin is slow to simultaneously introduce the required credible limits on the state with respect to both subnational authorities and private businesses. An important lesson that remains to be learnt by Russian politicians, it seems, is that the strength of one party does not have to be built at the expense of the other. The idea that there can be several independent and strong players in the public sector appears to be still foreign to Russia.

3

The Territorial and Administrative Structure of Local Governments

Designing the territorial and administrative structure of the public sector is an important first step in structuring the system of intergovernmental fiscal relations, since this organizational structure underpins the institutional set-up under which fiscal decentralization will be implemented and administered. In considering the territorial and administrative structure of Russia's subnational government level, we take into account a variety of legal, organizational, institutional, and administrative issues. The policy issues all have an impact on the sound design of a system of decentralized governance that provides local governments with enough autonomy to perform effectively; ensures local accountability; embodies the right incentives; operates with maximum efficiency by delivering local public services at a minimum cost; and satisfies the needs and preferences of citizens.

In the analysis of the territorial-administrative structure of a country's public sector, there is a clear need to look beyond the narrow issue of the organizational structure of government, which is limited to considering how many levels of subnational governments exist and how the different levels of government interact. For instance, we need to clearly distinguish, from the outset, between truly decentralizing decisionmaking powers (autonomy) to regional and local governments and using local governments for the deconcentration of functions from the federal and regional government levels. Russia will only realize the full benefits of the fiscal decentralization process when an appropriate, significant degree of budgetary autonomy is granted to the local government level, and when local authorities and managers are made accountable for results. Similarly, the benefits from decentralization will not be complete unless regional governments are also able to consolidate and exercise a significant degree of budgetary autonomy. Conversely, intensifying the level of deconcentrated federal government activities and functions at regional and local levels may lead to the effective dismantling of Russia's decentralization agenda, which some accuse President Putin of having been pursuing since he took office.

Underpinning the issue of subnational government autonomy and the nature of the territorial structure of government is the organization and operation of the subnational governments' civil service. In particular, the extent to which subnational governments have control over the employment decisions (for example, hiring, firing, organization, and promotion) of local government employees and whether regional and local government are able to determine their own salary structure are some of the most complex issues in public-sector management, but also some of the most important determinants of subnational autonomy. No fiscal decentralization is complete without some degree of administrative decentralization. The biggest item in this process is granting local governments authority to fire and hire their employees and to set salary scales independently of the central authorities. Because payroll is easily the largest item in local budgets, hiring and salary decisions may have strong impacts on the fiscal viability of local governments.

This chapter is organized as follows. The first section reviews the legal framework for the organization and operation of local governments from the beginning of the transition to the present day. It is followed by an overview of the current organizational structure of the local government sector. Next we take a look at the evolution of the institutional changes within the public sector during the transition in the Russian Federation. We then move on to a discussion

on the local governments' discretion to determine public pay. In conclusion we list current policy concerns survey proposed solutions and recommendations made by the Kozak Commission, and offer other options for reform. Three appendixes provide a review of the normative theory on and the international experience with government structure and an assessment of the need for local government civil service deregulation in the Russian Federation.

The Legal Framework for Local Governments

Politically autonomous local governments were introduced in the Russian Federation with the first law on local self-government in 1991.¹ The law introduced dual subordination of local executives both to the local council and to the higher-level executive bodies in the regions. Moreover, the law required direct election of the local executive head. The law also granted local governments formal autonomy in formulating local budgets. However, at the same time a minimum expenditure budget was required to be set at the previous year's level of expenditures, adjusted for new expenditure responsibilities imposed by higher-level governments. Local revenues were supposed to be fine-tuned (or "regulated") by higher-level government in order to provide sufficient resources to finance the minimum expenditure budget. The law established the scope of competences for each tier and type of local government. However, except for administrative autonomy and the authority to regulate land use and merchandise, the law did not enumerate specific functions of local government. Instead, the law mostly contained tautological clauses assigning local governments responsibilities for funding and maintaining the facilities they owned.

Besides an unclear assignment of functions, autonomy of local governments under the 1991 local government legislation was further hampered by the lack of taxing authority. Only the first (higher) tier of local government (subregion cities and rayons) was allowed to introduce local taxes. Furthermore, the 1991 *Law on the Basic Principles of Taxation* limited the tax instruments available to local government to land and property taxes, business licenses, and over 20 minor taxes, including many so-called nuisance taxes (such as a tax on dog tags).

The system of local governance was further transformed with the introduction of the 1993 Constitution. The Russian Constitution guarantees local communities the right to govern local affairs separately from the sovereignty of the state. However, the Constitution established a general framework for local self-governance without any references to the existing administrative structure. The implementation of the Constitutional right for local self-government is assigned to the joint responsibility of the federal and regional governments. Thus the Constitution set the stage for the federal government to enact a number of constitutional laws that further developed the framework for local governance. The most important of these laws included the *Law on the General Principles of Organization of Local Self-Government in the Russian Federation*, the *Law on the Constitutional Right to Local Government Elections*, and the *Law on the Financial Foundations of Local Self-Government in the Russian Federation*.

The *Law on the General Principles of Organization of Local Self-Government in the Russian Federation* establishes the concept of a municipal settlement as the basis institutional unit for local self-government, where the municipal settlement is defined as *any* inhabited territory having three principal attributes: municipal property, a local budget, and elected bodies of local self-government. Regional governments were supposed to establish a system of local self-government on their territory along the federal guidelines by passing regional legislation on municipal boundaries, municipal elections, and so on. In practice, regions had several options for superimposing the institution of the municipal settlement over the existing

1. Federal Law No. 1550-1 on *Local Self-Government in the RSFSR* [Russian Soviet Federated Socialist Republic] (July 6, 1991).

hierarchy of local governments.² The majority of the regional governments in Russia established municipal settlements exclusively at the first (higher) tier of local governments (that is, at the level of subregion cities and rayons). In 15 regions, municipal settlements were established only at the second (lower) local government tier (towns, townships, and rural districts). In the rest of the regions, municipal settlements were established at both the first and second tiers of local government, resulting in a system of local government in which some municipal settlements are subordinate to others.

The Organizational Structure of Subnational Governments

A comprehensive overview of the structure of local government currently in place in the Russian Federation is provided by Kourliandskaia, Nikolayenko, and Golovanova (2002). Based on the foundations of local government contained in the Russian Constitution, already briefly discussed above, and the *Law on the Principles of Local Self-Government*, a local government (or municipality) is defined as any populated territory (city, town, township, a collection of these on a single territory, or any part thereof) to which the following four conditions apply:

- It is self-governed.
- It has municipal property.
- It has a budget.
- It has elected bodies of local self-government.

At present, there are about 29,500 units of local administration below the regional level in Russia (see table 3.1). Compared to the international experience with government structure, the average settlement size in Russia of 5,400 is at the lower end of the size range of basic authority observed in other countries, ranging from 1,560 persons in France to 122,000 persons in the United Kingdom (appendix 3.2). Moreover, as recorded by Kourliandskaia, Nikolayenko, and Golovanova (2002), of these local administrations there are only 12,261 that are officially registered as municipal entities; only 11,691 have elected representative authorities; only 11,209 municipalities are endowed with municipal property; and only 4,500 have fully independent budgets.

In conformity with the Constitution, the *Law on the General Principles of Local Self-Government* of 1995 allows for regional governments to determine the approach by which local self-government is instituted, meaning that regions could permit local governments to be established at the rayon level as well as at the level of subrayon towns or villages, irrespective of their population. A complicating factor is that the law grants a uniform legal status to all local government entities. This means that all local governments—be they rayons, towns, or townships of rural districts—enjoy the same institutional and administrative rights. This has been interpreted to mean that the law does not allow “subordination” of one municipality to another. Kourliandskaia, Nikolayenko, and

Table 3.1 *The Inherited Administrative Structure of Territorial Division*

<i>Tier</i>	<i>Administrative unit</i>	<i>Number of units</i>	<i>Average population</i>
1	Central government	1	145.5 million
2	Federal districts	7	20.8 million
3	Regional-level units	89	1.6 million
4	Rayon-level units	2,513	57,900
5	Township-level units	27,120	5,400

Note: As of January 1, 2000.

Source: Goskomstat (2001).

2. For an overview of these issues, see Kourliandskaia, Nikolayenko, and Golovanova (2002).

Golovanova (2002) thus conclude that pursuant to the *Law on the General Principles of Local Self-Government*, the entire territory of the Russian Federation should be divided into non-overlapping jurisdictions (territories) of municipalities or local self-government entities.

The current practices across Russia's regions in defining local government structure do not always conform to the legislative requirements. While Russia historically has had a very rigid administrative hierarchy in which administrations of smaller territories were subordinate to higher-level administrations (see chapter 2), municipalities that have come into existence in recent years have formed at different tiers of the former administrative and territorial hierarchy. All of these factors help provide a remarkably diverse scene. Kourliandskaia, Nikolayenko, and Golovanova (2002) distinguish the following three basic types of local public administration currently in operation in different regions of Russia:

Type I—One level of municipal bodies (local governments) at the level of large cities and rayons. In this form of local government structure, only one tier of local bodies of self-government is formed in the region, so that the regional government has to relate only to one set of local governments. The region can establish local governments at the subrayon level (denoted as a Type I-A local government structure in box 3.1) or on the basis of preexisting boundaries of the large cities and rayons in the traditional administrative-territorial division (denoted as a Type I-B local government structure). Municipalities enjoy the fiscal rights determined by federal legislation. Their fiscal relations with "state power" (the federal and regional government apparatus), such as the determination of intergovernmental transfers and tax sharing, are designed by the regional authorities.

In a number of regions the single-tier regional administration coexists with municipalities formed on the basis of subrayon cities and rural districts (a Type I-C local government structure). To the extent that rayon-level governments are not relied upon, regional authorities have to deal with hundreds of subrayon municipalities formed on the basis of small villages.

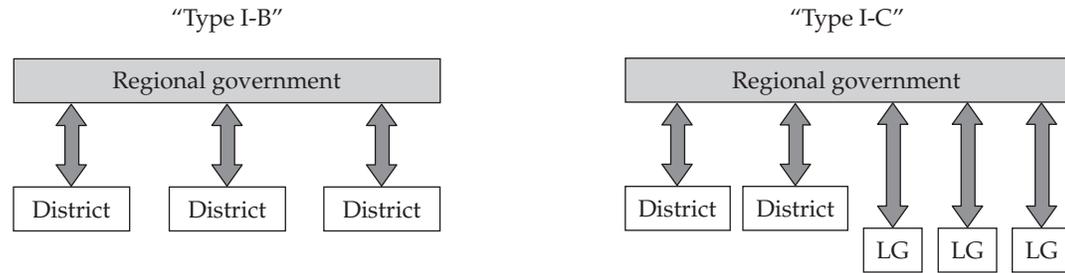
Type II—One deconcentrated tier of regional government at the local level combined with one level of municipal bodies (local governments). This local structure is similar to the previous one except that the authority to set local taxes is vested in rayon-level territorial branches (that is, deconcentrated units) of the regional administration that have no elected bodies. Budgets of municipalities are incorporated in the regional budget in the form of expenditure plans. In this approach, municipalities (especially those formed at below-rayon level) are deprived of some budgetary rights and powers established for them by federal legislation.

Type III—Two levels of local self-government where some local governments are subordinate to other local governments. In this form of local government structure, the first level of local government includes big cities and rayons; the second level includes subrayon cities and other incorporated populated areas within cities or rayons. Local residents elect both the first-level and the second-level governments. Local governments of the first level engage in direct financial relations with the regional authorities. Financial relations between localities of the second level and regional authorities are indirect and go through local governments of the first level. Local governments of the first level are responsible for the distribution of regional grants to localities of the second level and for splitting the sharing rates of regional-federal shared taxes assigned by the regional administration to the local level. As such, first-tier local governments perform the functions of bodies of State power assigned to the regional governments.

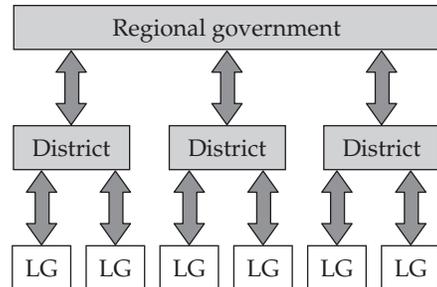
In cases when second-tier (in other words, lower level) municipalities cover the entire territory of a first-tier municipality, the rayon-level local government is usually only responsible for the maintenance of social infrastructure facilities that provide services to the entire population of the area. Local taxes in such cases are usually established by the second-tier municipalities, but sometimes these smaller municipalities delegate their taxing authority to the rayon-level municipality, in which case the rayon usually sets a uniform tax rate for all municipalities located within its boundaries. If second-tier municipalities represent individual inclusions in the territory of the first-tier locality, the latter is generally responsible for the performance of

Box 3.1 *The Various Structures of Local Government in the Russian Federation*

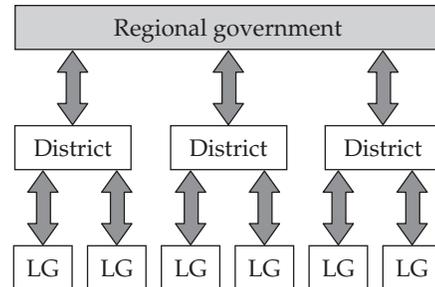
These are three examples of a **Type I** local government structure, where the regional (oblast) government deals with one level of local government, either at the district (rayon) level or the local (town or village) government (that is, subrayon) level. Note that different types of local governments should still be considered one local government level if they are not hierarchically structured.



This is an example of a **Type II** local government structure, where the district level is a deconcentrated tier of the regional government and local governments are deprived of some of their budgetary rights.



This is an example of a **Type III** local government structure, where two levels of local government exist and local governments are hierarchically subordinate to district-level governments.



LG. Local government.

functions of local self-government in the territory outside the jurisdiction of second-tier municipalities. Local taxes in this case are established by local governments of both levels.

As noted by Kourliandskaia, Nikolayenko, and Golovanova (2002), the forms of local self-government organization summarized in box 3.1 can make up different combinations in different subjects of the Russian Federation: for example, of 295 municipalities of Tyumen Oblast, 4 have been formed at the large-city level and 2 at the rayon level, with the rest set up at the level of small townships below rayons. In Vladimir Oblast, 18 municipalities set up at the rayon level contain lower-level local governments while the other 7 municipal entities have a one-tier type of organization. The second-tier municipalities in Vladimir Oblast are individual populated areas that have opted for local autonomy. However, regional authorities are unwilling to recognize them as full-fledged participants in intergovernmental fiscal relations, choosing to deal with them through the first-tier bodies of local government.

The Evolution of the Public Sector during the Transition

The organization of the public sector in Russia represents a special interest for the work in this book from a number of angles. First, given that both the nature and the functions of the State have drastically changed since the fall of the Communist regime, it has to be seen whether the corresponding necessary structural adjustments actually have taken place in the structure of the public sector. Second, we are interested in seeing whether newly acquired responsibilities of subnational governments have been matched with adequate capacity to deliver public services. This capacity is determined by existing public institutions and the personnel employed by them. Finally, no fiscal decentralization can be considered complete without some degree of administrative decentralization in firing and hiring employees and setting salary scales independently of the central authorities.

Public Entities

Modern Russia's public sector has its roots in the Soviet government system. During Soviet times, essentially all the economy was composed of one single (public) sector. Nominal distinctions were drawn among the budgetary sphere, the state enterprise sector, and the emerging private (or collective) sector. But, in reality, these were "only hazy notions of which and how many employees were carrying out what jobs in government and who paid them" (Nunberg 1999).

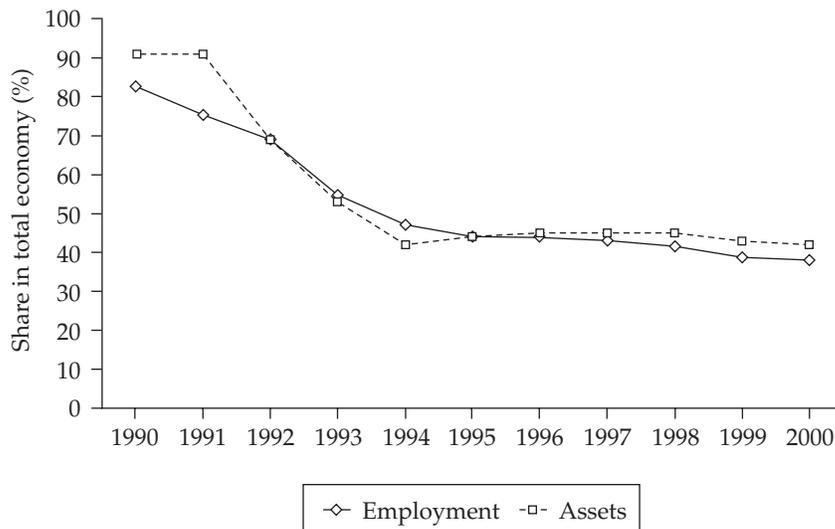
The "budgetary sphere" was formally defined to include the central government, subnational governments, and "nonproductive" and "budget-supported" institutions and social facilities (such as healthcare, education, science and culture, waste-management, and so on). The terminology of the "budget sphere" continues to be used today in defining the scope of public sector activities in the Russian Federation.

The definition of "budget sphere" approximately corresponds to the notion of general government as used in many countries around the world today, following the International Monetary Fund (IMF) standard classification. However, fully equating these two concepts clearly would lead to two classical types of errors. On the one hand, the Soviet budgetary sphere used to include central directorates of industrial agglomerations. On the other hand, centrally planned enterprises were responsible for the provision of many basic goods and services. These enterprise facilities were not accounted for in the budgetary sphere even if they were partially funded from the budget.

Before being privatized state enterprises maintained significant social assets—housing, day-care, hospitals, and recreation facilities. In addition enterprises were partially responsible for maintaining school buildings through a scheme of enterprise-school partnership. In the mid-1990s, privatization was accompanied by the process of divestiture, meaning the transfer of social assets and the responsibility for their financing to municipalities, or simply the closing down of those facilities. Thus the core activities of the enterprises were moved out of the public sector while essentially public services were made formally part of the public sector.

The dramatic reduction of the public sector between 1991 and 1995 is shown in figure 3.1. The share of public assets in total national wealth decreased from 91 percent to 42 percent. At the same time the share of the labor force employed in the public sector decreased from 83 percent to 40 percent. However, these numbers might provide an overestimate of the actual scale of the privatization. For example, the reclassification from public to private was based on the legal transformation of a state enterprise into a joint-stock company. Often, however, the state would still remain a major shareholder of the privatized company.

Although the public-private dichotomy may be a bit misleading in terms of the extent of state involvement in the economy, we have no choice but to assume that the entities that have been reclassified as private are those not engaged in the provision of public services. Thus the post-privatization public sector presumably consists of three subsectors: the public administration

Figure 3.1 *The Size of the Public Sector*

Source: Authors.

(legislative and executive bodies of government), state (or municipal) “institutions,” and state (or municipal) unitary enterprises.

The law defines an “institution” as a not-for-profit organization established and financed, fully or partly, by its owner.³ Assets are transferred to an institution on the basis of operational management. The bodies of state power (the federal and regional governments) and the bodies of local government establish state and municipal institutions correspondingly. Thus, institutions are public entities—such as schools and polyclinics and the like—that are funded primarily through budget appropriations. Usually, the legal form of state (or municipal) institution is used for the administration of public facilities providing public services such as healthcare, education, culture, and others. Note, however, that theoretically, healthcare institutions can receive a larger portion of their funding from the Health Insurance Fund than from the budget of the government-owner (see box 3.2).

The *Civil Code* defines a state (or municipal) unitary enterprise as a commercial legal entity whose property is transferred from the state (or municipality) under the right of either operational management or economic management. The latter legal regime provides that state-owned enterprises “do not have the right to sell, rent out, mortgage, contribute as an investment into the charter capital of another company or partnership, or dispose of in any other way, the immovable property which belongs to them, without the consent of the owner. The legal form of a unitary enterprise is mostly used for public facilities that recover a larger share of their costs from sales (for example, public transit, utilities, and so on).

Finally, it should be mentioned that there are a few instances of public service delivery through nongovernmental organizations (NGOs). Several regions and cities have introduced special amendments to their local legislations to allow for contracting out municipal services.⁴ Some of these initiatives, sponsored by local NGO communities and international donors, aim at enabling NGO participation in competitive bidding to produce social services funded by municipal grants. However, this practice is very limited and most likely it involves old NGOs that used to be on the government payroll during the Soviet period (for instance, the association of people with disabilities).

3. The federal *Law on Not-for-Profit Organizations* of 1996.

4. To our knowledge these include the City of Moscow, Voronezh Oblast, and the Republic of Buryatia.

Box 3.2 *Russia's Healthcare System*

The Constitution assigns the responsibility for healthcare provision jointly to the federal and regional levels of government. Thus federal legislation establishes a general framework for healthcare regulation and financing while the implementation is left to the discretion of regional legislatures. Hence, the resulting healthcare systems present a blend of the concepts envisioned by the federal authorities and the actual regional government implementation aiming at achieving their desired outcomes (most often status quo) within the federal framework.

The federal healthcare legislation originated in the experiments of the late 1980s that introduced new economic mechanisms to the sector. Throughout the Soviet period, healthcare facilities were funded based on their capacity and the moneys were often channeled through the enterprises that formally maintained these facilities for the benefit of their workers. The reforms aimed at extending the regulatory powers of local authorities and introducing some self-management for the healthcare facilities. Thus, Territorial Medical Associations (TMA) were established on the bases of polyclinics to receive financing based on the number of registered residents and to make payments to hospitals for treatment received by the covered population.

The current federal framework for healthcare provision was established with the *Health Insurance Law* and the *Protection of People's Health Law*. These laws establish universal access to healthcare and make the local level of government responsible for ensuring accessibility of healthcare services. The federal government guarantees free provision of services covered by the basic program of Compulsory Health Insurance (the coverage can be further extended by regional authorities). Certain types of medical care are to be funded directly from federal and regional budgets.

The Compulsory Health Insurance program is carried out through the system comprising the Federal Health Insurance Fund, Regional Health Insurance Funds, Health Insurance Organizations, and licensed healthcare facilities. Healthcare services are provided at the licensed facilities and paid for by Health Insurance Organizations on fixed tariffs. However, because the tariffs are set at a very low level, most of the participating facilities are those owned and subsidized by local and regional governments. Health Insurance Organizations are independent legal entities receiving a fixed amount from the Regional Health Insurance Fund for each insured person. As Health Insurance Organizations have to be licensed by the Regional Health Insurance Funds, they are dominated by those established by local and regional authorities.

Regional Health Insurance Funds are independent public nonprofit credit institutions accountable to regional authorities. Regional Health Insurance Funds have three sources of funding: contribution from employers for the working population (3.2 percent of payroll); contributions from local governments for the nonworking population (set by regional authorities); and equalization transfers from the Federal Health Insurance Fund (in turn formed with employers' contribution of 0.4 percent of payroll). The low level of employer contributions does not cover the costs of the basic Compulsory Health Insurance program. Consequently local authorities have to make up the difference either by covering losses of healthcare facilities that they own or by making larger contributions to the Regional Health Insurance Fund for the nonworking population.

Source: Potapchik (2002).

Employment within the Budgetary Sphere

The budgetary sphere roughly corresponds with the notion of general government, and those employed within the budgetary sphere are responsible for the provision of core public services. In Russia, these are personnel employed in the government bodies and state (municipal) institutions. Because these personnel establishments are funded from the government budget, theoretically the total number of personnel can be obtained by using the "Personnel and Client Groups" Form, which is used to estimate labor costs during budget planning. In contrast, state unitary enterprises enjoy the right of economic self-management, and they receive government funds, if any, in the form of production subsidies, rather than itemized appropriations. Hence, data on employment in unitary enterprises are not systematically collected. However, we are less interested in the latter portion of public employment because the types of services they provide (such as utilities) tend to belong to the private sector in many other countries. Moreover,

recently we witnessed the first cases in which municipalities made concessions for municipal utilities' provision to private companies in Russia.

Even though the data on employment in the budgetary sphere are obviously maintained by the respective governments as part of the budget process, there are no readily available consolidated data for the nation as a whole. The official statistics bureau (*Goskomstat*) records employment only by branch of economy (for instance, education, healthcare, and so on) but does not distinguish between public and private employment within each sector. The official statistics explicitly report public employment only for public administration (that is, government bodies but not service facilities). For the budgetary sphere employment other than in the public administration, we rely on the study by Nunberg (1999) for 1990–92 and on our own estimates based on a partial survey of subnational government in 2000.⁵

According to Nunberg (1999), in 1992, Russian general government employment—estimated at 21.3 percent of the labor force⁶—was higher than the lower end of Organisation for Economic Co-operation and Development (OECD) countries (6.1 percent in Japan, 13.8 percent in Italy and Spain, and 14.9 percent in the United Kingdom) but still below Denmark (27.4 percent), Finland (28.6 percent), Norway (28.7 percent), and Sweden (35.7 percent). As of 1992, the bulk of budgetary sphere employment was concentrated in two mega-sectors: education and healthcare (including physical culture and social security). Employment in these two sectors amounted to 10.6 million (15 percent of total civilian employment in 1992). In comparison with those in industrialized countries, the health sector was relatively small while the education sector was proportionally large. Healthcare in Russia represented only 5.9 percent of civilian employment (versus 8.5 percent in the United Kingdom and 10.4 percent in the United States) while the education sector accounted for 9.0 percent of civilian employment (versus 6.7 percent in the United Kingdom and 7.4 percent in the United States). This initial misalignment combined with an aging population foreshadowed a likely structural adjustment within the “budgetary sphere.”

The responsibility for these two sectors (education and health) is mostly assigned to subnational levels. In 1992, regional and local governments accounted for 66 percent of total expenditures on education and this figure had increased to 86 percent by 1997 (Martinez-Vazquez and Boex 2001). Subnational government share in total expenditures on healthcare has been about 88 percent throughout the transition years.

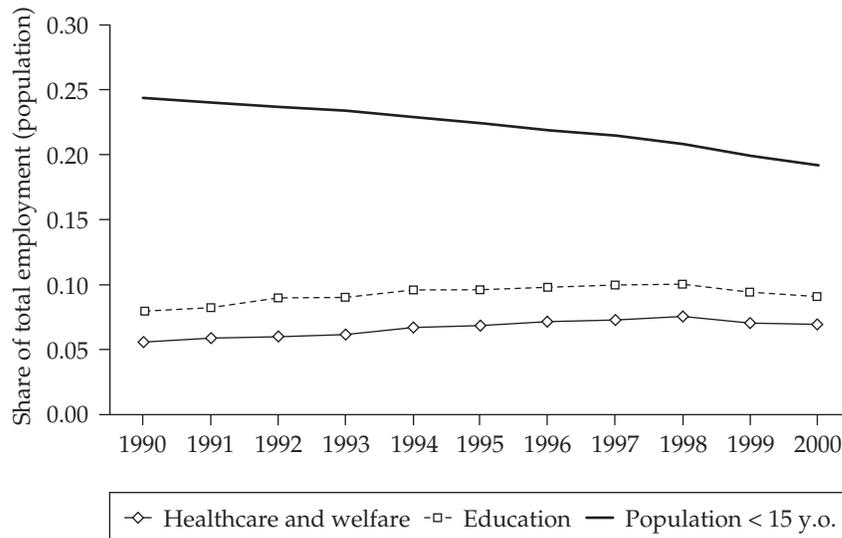
As we will discuss in chapter 6 (on intergovernmental fiscal transfers), because of vertical fiscal imbalances, many local and regional governments depend heavily on intergovernmental transfers. Unfortunately, for a while at the federal-regional level, the design of the grant distribution has not provided the proper incentives for subnational governments to carry out the necessary structural adjustments—often still the case at the regional-local level. The amount of the transfers is determined frequently by the capacity of existing public facilities, and therefore the rationalization of existing capacity and the closing of redundant infrastructure would lead to cuts in transfers.

These aforementioned perverse incentives might explain the irrational development of the employment distribution between healthcare and education during the 1990s. As figure 3.2 shows, the share of the population below the working age has been steadily decreasing—from 24.4 percent in 1990 to 19.3 percent in 2000. At the same time, employment in the education sector had been increasing until 1998.⁷ Moreover, employment in the much-needed healthcare sector had been growing at the same pace as in education. Thus in 2000 the “misalignments”

5. The survey was conducted by the Center for Fiscal Policy (Moscow) through requesting the “Personnel and Client Groups” Form from individual regional governments. For every region the form also covered budget sphere employment in constituent localities. The number of responding regions totaled 47 out of a total 89, with about 1,200 local governments covered by the survey.

6. Including central and subnational governments but not state enterprises.

7. These figures include both public and private facilities of healthcare and education. However, because private provision of these two service types is minuscule, we attribute the observed trend to the public sector.

Figure 3.2 *Structural Adjustments in the Public Sector*

Source: Authors.

between the two mega-sectors were roughly the same as in 1990. In spite of some adjustment, which took place during 1999–2000, we still observe overemployment in the education sector and underemployment in the healthcare sector by comparison to the profiles of other industrialized countries.

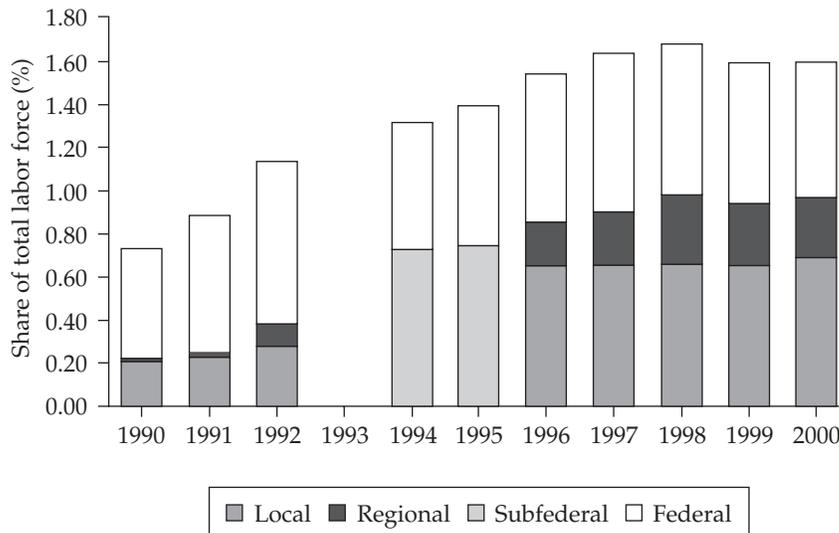
Over the 1990s, Russians' health status has deteriorated sharply, which has been manifested in the development of preventable disease epidemics, such as tuberculosis and HIV/AIDS (Vinokur and others 2001). The unprecedented decline in Russians' life expectancy between 1987 and 1994, followed by an unexpected recovery by 1998, calls for in-depth research. Existing studies suggest that the decline in life expectancy, showing large variations among Russian regions, cannot be attributed simply to impoverishment, as the wealthiest regions experienced the greatest drops (Waldberg and others 1998). The evidence presented in this section suggests that Russia's poor public health outcomes might have been caused in part by the misallocation of scarce fiscal resources at the subnational level.

Public Administration

During the 1990–92 period, governments' employment as a whole contracted by 5.1 percent. However, over the same period, public administration staff increased by 24 percent (figure 3.3). Changes in government administration during 1990–92 were mainly due to the dramatic shifts in sovereign responsibilities that occurred with the dissolution of the Union of Soviet Socialist Republics (USSR) and the creation of the Russian Federation. Certain sovereign functions once reserved to the Soviet Union including those performed by the Ministry of Nuclear Power, the Ministry of Railroad Communication, and the Informatics Committee, were assumed by the Russian government. As a result, the number of federal personnel increased from 383,000 to 536,000 between 1990 and 1992.

Federal personnel were, and remain, overwhelmingly deconcentrated, however (see figure 3.4). In 1992, 94 percent of federal personnel worked in the territorial offices of federal agencies outside Moscow. Deconcentrated federal employment at the regional level numbered about 144,000. Federal employees were most numerous at the subregional level, numbering 358,000. At the same time, regional and local governments including legislative branches employed about 271,000 in 1992, substantially up from 168,000 just two years earlier. This 61 percent jump

Figure 3.3 Public Administration Staff



Note: 1993 values are not available.

Source: Authors.

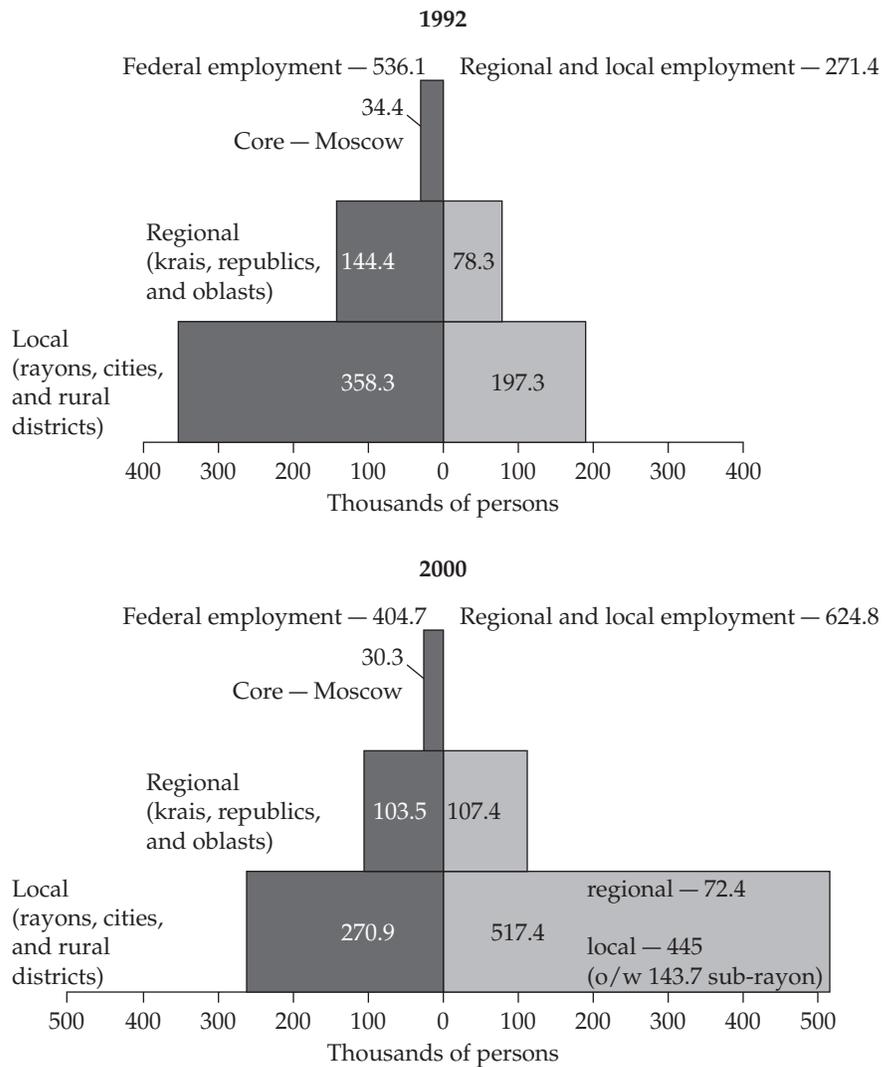
reflects the increased importance of the subnational governments in Russia in the context of significant decentralization of governmental functions. However, as of 1992, federal employees still outnumbered subnational personnel by two to one.

The proportions between federal and subnational employment in the public administration reversed after the 1993 constitutional crisis. With the new government system, the heads of the executive branch of subnational governments started to be elected directly,⁸ and thus they gained some autonomy from their respective legislatures and also from the federal and upper-level government authorities. Some ethnic republics of Russia had their executive head directly elected even before 1993, as they did not want their governance to be anything less than that of sovereign Russia. In 1994, the ratio of federal to subnational public administration employment was 0.74. This ratio recovered to 0.82 by 1997 but it has been decreasing ever since.

Excepting the turmoil of 1993 and the aftermath of the 1998 financial crisis, the size of the public administration as a whole has been growing at a steady rate. As can be seen in figure 3.4, the 2000 structure of the public administration presents some of the features existing in 1992, but at the same time reveals a few radical changes. The Soviet legacy is mostly obvious in the vertical distribution of federal employees. Although the territorial portion of the federal public administration decreased from 94 percent to 75 percent, federal government remains highly deconcentrated. While the territorial staff was cut by 25 percent, the Moscow core employment, rather than increasing, actually also decreased by 10 percent. One potential consequence is that the small core at the center is likely to preclude the realization of economies of scale. Moreover, the lack of coordination and logistic support from the Moscow headquarters makes federal employees in the field dependent on the support of local authorities. This blurs the delineation of powers between the levels of government and leads to the *de facto* "dual subordination" over the federal employees in the field.⁹

8. The federal *Law on the Appointment and Dismissal of Heads of Krai, Oblast, Autonomous Oblast, Autonomous Okrug, Rayon, Town, Urban Rayon, Settlement, and Rural Administrations* of 1993.

9. The phenomenon of *de facto* dual subordination has had a significant impact on the ability of federal government agencies to implement their own policies, which may be in relative conflict with the objectives of subnational governments. This conflict has been most visible in the area of tax administration and enforcement. See Martinez-Vazquez and Boex (2001).

Figure 3.4 *The Vertical Distribution of Federal and Subnational Public Administration Staff*

Source: Updated and modified from Nunberg (1999) and Goskomstat (2001).

Most changes in the structure of the public administration involved employees of regional and local governments. Over the 1990s, regional public administration staff increased by a factor of 18 and outnumbered federal employees located in regional centers. Only about 40 percent of regional public administration is deconcentrated to the city-rayon level. Thus at the rayon-city level, federal employees still outnumber regional staff by a factor of four. Nevertheless, local government employees are the most numerous at the city-rayon level and account for the largest share of total public administration in the country, with the subrayon portion accounting for 30 percent of local public administration staff.

The Evolution of Public Payroll of Subnational Governments

In 1992, by comparison to most other countries in the world, Russia had the lowest general government (excluding the military) wage bill in relation to the GDP—5.2 percent. This must be explained by the delivery of many public services by enterprises outside the budgetary sphere and the relative importance of non-wage benefits for public servants, which were accounted for in other public-administration expenditure subcategories or sometimes not considered as public

administration spending at all. For most of the Soviet period, conditions of suppressed inflation were associated with minimal changes in personal incomes (Nunberg 1999). Thus, shifts in salaries and pensions were not decisive factors in public personnel management. Much more important than salaries was privileged access to goods and services, the volume and quality of which were strictly regulated.

Evolution of the Unified Pay Scale

At the end of the 1980s, under Gorbachev's economic reforms, inflation accelerated, and for the first time the curtailment of in-kind privileges for middle-ranking public servants occurred. This meant that base salaries began to attain greater importance relative to these employees' total income. Inflation spurred the state to embark on a policy of public-service salary increases and to install something similar to a pay and grading system. In September 1989, the USSR Council of Ministers passed Regulation No. 803, "On the Improvement of the Remuneration System for the Employees of the State Administration Apparatus," which detailed conditions for the employment, promotion, and remuneration of public administration employees at different government levels. However, the new public administration regulation along with the existing remuneration norms for all kind of public entities fell prey to the disintegration of the Soviet Union. As a result of political decentralization, authorities of Soviet Republics launched efforts to secure themselves increased compensation benefits. In addition, the disintegration of the pay system quickly trickled down to the regional and local levels. These efforts contributed to the tendency toward erosion of centrally established norms and rules for remunerating employees of subnational public entities.

It was in this climate that the RSFSR Council of Ministers issued Resolution 436 on August 19, 1991, establishing a unified salary scale for all public employees except those in the public administration. This unified scale represented an attempt to simplify centralized norms. It was based on 10 basic grades with differentials of 10–20 percent among them. The grading structure was based on the existing titles, which were often driven by political criteria. As a result, although grades were officially determined by function, posts were often ranked in relation to the number of the Communist Party members in the organizational unit (Nunberg 1999). Subsequently the 1992 price liberalization followed by hyperinflation required frequent indexation of the public wages.¹⁰ As some of these raises covered only selected areas of the public service, the initial pay scale became distorted very soon.

In October 1992, the federal government introduced a new unified salary scale.¹¹ The new scale aimed to restore work incentives by decompressing public wages. It had 18 grades and appeared to be excessively complex, perpetuating problems of the old system such as too many increments and automatic increases (Nunberg 1999). Moreover, it proposed to decompress salaries to what was probably an untenable level, creating an excessively hierarchical structure. The scale comprised the basic rate (for the first grade) and 17 adjustment coefficients, by which the basic rate is multiplied to derive rates for higher grades. The coefficient for grade 12 was 5.1, while for 18 (the highest grade) it was 10.07.

Current Pay Determination System

In the current system, pay determination differs between civil servants and other budgetary sphere employees. The status of civil servant is given to an employee filling a position within a

10. Some of the initial increases were by a factor of 1.9 in December 1991 (Presidential Decree No. 211); by a factor of 1.45 in February 1992 (President's Order No. 28-RP); and by a factor of 1.5 in September 1992 (Presidential Decree No. 895).

11. Government Resolution No. 785 of October 14, 1992 on the Differentiation of Public Pay Based on the Unified Salary Scale.

special registry of posts. These posts involve exercising government authority. Thus, currently, the distinction is based on the position rather than the type of employer (administration versus service delivery) as was done before. Furthermore, the civil-servant status varies with the level of government. Civil servants at the federal and regional levels are called "state servants," while at the municipal level the title is "municipal servant." The distinction arises from the constitutional separation of local self-government from the sovereignty of the state.

The basic legal underpinnings for the civil service at the federal and regional levels are established in the 1995 federal *Law on the Principles of State Service*.¹² The remuneration for each grade has to be set by the federal or regional governments within the framework of this law. This piece of legislation has been criticized for both being too detailed and compromising on certain points of privilege and remuneration, which should be left to subsequent regulations, and too loose and noncommittal on certain crucial definitions of institutional codes and division of powers (Nunberg 1999). The law contains, for example, detailed provisions on grading, promotion, and recruitment procedures as well as pension provisions for civil servants. Locking in such detailed policies could prove financially costly if implemented. The law also enshrines some questionable practices stemming from past procedures, such as increments based on length of service, which could undermine the possibilities of introducing performance-based incentives into Russia's new civil service in the future. Many of the provisions on remuneration, moreover, appear to be focused on rewarding currently serving public employees rather than establishing transparent, generic standards for the future civil service.

The legal regulation of the municipal civil service has to be prescribed in the charter of each municipality conforming to the requirements of the regional legislation and the federal *Law on the Principles of Municipal Service*.¹³ These regulations should cover qualification requirements, municipal employee status, terms and conditions of municipal service, and administration of municipal service. Regional laws usually establish limitations on the level of wages for municipal servants. Since the majority of municipalities are subsidized by the regional budget, the regional governments often try to prevent municipal employees from earning more than their regional counterparts (Kourliandskaia, Nikolayenko, and Golovanova 2002).

Unlike civil servants, individuals employed in the budgetary sphere (for instance, teachers, public health workers, university professors, and so on) are covered by the unified pay scale that continues to be governed by the 1992 government resolution, subject to indexations decreed via subsequent resolutions.¹⁴ The adjustment coefficients were augmented in June 1994, replaced with a more compressed scale in February 1995, and restored to their original form in August 1995. In 1998, the unified salary scale was restated in a federal law,¹⁵ establishing that the basic rate should be set in a federal law while the adjustment coefficients for other grades should be determined by the federal government after consultation with national labor unions and associations of employers. This law was first vetoed by the President; however, the veto was overruled by a majority of the State Duma. The law was prompted by the federal government's failure to restore the purchasing power of public wages after the multifold devaluation of the ruble in August 1998.

The consultations with labor and employer associations are supervised by the federal Ministry of Labor and Social Welfare. This ministry also produces technical documentation for the revision of pay, grading, and benefit policies, along with drafts of legislation detailing public pay proposals. It is also officially empowered to issue decrees on remuneration matters.

12. Federal Law # 119-FZ of July 31, 1995.

13. Federal Law # 8-FZ of December 17, 1997.

14. Government Resolutions No. 759 of June 30, 1994; No. 189 of February 27, 1995; No. 823 of August 24, 1995; No. 585 of May 13, 1997; No. 309 of March 18, 1999.

15. Federal Law (# 22-FZ of December 2, 1998) on *Arranging the Pay of Employees of the Budget Sphere Entities*.

In practice, the remuneration of public employees continues to be highly decentralized. Indeed, within the central government, as well as at the local and regional government levels, considerable discretion is given to heads of entities and their department chiefs with regard to the awarding of cash bonuses and allowances and benefits. Thus, apart from the regular wages, the public sector remuneration in Russia is further determined by an assortment of prescribed monetary increments that augment the basic wage. These include allowances (up to 8 percent of total salary), bonuses (3 months' equivalent salary), and "material assistance" (2 months' equivalent salary). Together these can add up to more than 40 percent of the total remuneration. The system operates on a monthly basis and the distribution is determined by the chief of each government agency, often in consultation with the personnel department and the trade union committee. In addition, personal bonuses, awarded on unspecified criteria, can comprise 15–20 percent of the basic wage and are incorporated permanently into the salary. Some entities also exercise managerial flexibility by using discretionary funds derived from nonbudgetary sources, such as reserves or commercial operations, to top up salaries and benefits for employees.

The coverage of the national pay scale for subnational governments has varied over time. The initial decree of 1992 established that the "unified pay scale is mandatory for all entities receiving budget funding," thus including employees within the budgetary sphere at the regional and local governments. However, the 1994 raise was decreed only for "the budgetary sphere entities funded by the federal government." This decree explicitly stated that for subnational governments the raise was only recommended and referred to the provision of the 1993 Presidential Decree,¹⁶ which had established that unfunded federal orders were not binding for subnational governments. The salary raises of 1994, 1995, and 1997 again were decreed only for federal entities and established that subnational governments give raises "independently and at their own expense." With the introduction of the 1998 federal law, however, the unified pay scale became strictly binding for all "budget sphere entities." For example, based on this law, a regional union of teachers sued a regional government in the Supreme Court for postponing the pay raise for six months.¹⁷ The court ruled against the regional government on the grounds of superiority of federal laws over any subnational decision.

Since the 1999 law, salaries paid by subnational governments to budget sphere employees cannot be below the federal statutory level. Moreover, federal legislation entitles employees to a special allowance if the entity is located in the Far North or an equivalent territory. Local authorities can pay local allowances to employees of municipal budgetary institutions (primarily schools, outpatient clinics, and hospitals) in addition to the basic pay, but rarely exercise this discretion owing to lack of financial resources.

The federal government has often had to act when subnational governments accumulated huge arrears in wages to public employees. Thus during the 1996 presidential campaign, the federal government provided loans to regional governments to clear wage arrears in subnational budgetary entities (OECD 2000, p. 140). Complete clearance of subnational wage arrears occurred only in 1997, when the volume of federal loans exceeded 40 percent of the explicit federal grants (see chapter 7 on subnational debt). Not surprisingly, given the moral hazard involved in this type of federal rescue but also in combination with general economic conditions, by 1998 the volume of wage arrears accumulated by subnational governments was approaching the 1996 level again. The federal government had to roll its loans over to the regions while introducing certain conditionalities. Among these conditions were one-to-one matching of regional funds allocated for clearing arrears and dedicating a fixed share of regional budget to wage payments.

16. Presidential Decree No. 2268 on Forming the Republican Budget and Relations with Regional Governments in 1994.

17. Civil Case # 99pv-200pr: Krasnoyarsk Krai Union of Public Education and Science Workers vs. Krasnoyarsk Krai Administration.

Pay Variation

At the beginning of the transition, education and healthcare accounted, respectively, for 33 and 22 percent of the budget sphere wage bill (Nunberg 1999). In 1992, different regions showed roughly similar patterns of public wage expenditures, with education and healthcare accounting for about two-thirds of the subnational wage bill. In the 2000 sample of subnational governments, education on average accounts for about 50 percent of the local government wage bill and 19 percent of the regional government wage bill (table 3.2). However, healthcare accounts for only about 15 percent of the local government wage bill and 20 percent of the regional government wage bill. Note, however, that a significant portion of the healthcare wage bill is funded with receipts from the Health Insurance Organizations. Public administration accounts for about 19 percent of the local government wage bill and 25 percent of the regional government wage bill.

In 1992, average wages in the budgetary sphere showed significant variations of up to a factor of three. This variation was roughly consistent with the differences in the cost of living across Russia's national territory. In the 2000 sample of subnational governments, there is significant variation in the average public wage across jurisdictions, at both at the regional and local levels. At the regional level the standard deviation ranges from 47 percent of the mean wage in the education sector to 72 percent in law enforcement. At the local level the standard deviation ranges from 39 percent of the mean wage in the education sector to 50 percent in public administration. As expected, public administration employees have the highest average wage both at the regional and local levels. Adjusting by the price level slightly reduces the variation in all branches except law and order (see the last column of table 3.2).

In summary, pay determination differs between public-administration and service-delivery personnel. Full discretion given to each jurisdiction over remuneration of its civil servants results in considerable pay variation both among levels of government and among government units of the same level. The unified pay scale for service delivery personnel, which has been strictly binding for subnational governments since 1998, reduces pay variability in this sector. However, even in service delivery, average wage does vary from one government to another owing to the discretion in awarding supplements and bonuses, which can amount to 40 percent of the basic wage. The unified pay scale thus serves more like a minimum standard of service delivery. Subnational governments have to follow each national pay raise, which is often accompanied by a targeted loan from the federal government followed by additional loans if wage arrears continue to build up.

Table 3.2 *Average Wages of Regional and Local Public Employees, 2000*

	<i>Annual wage, thou. RUR</i>			<i>Ratio to the subsistence level</i>		
	<i>Mean</i>	<i>Median</i>	<i>C.V.</i>	<i>Mean</i>	<i>Median</i>	<i>C.V.</i>
<i>Regional governments</i>						
Total budgetary sphere	16.730	12.497	0.81	1.74	1.51	0.76
Public administration	50.768	45.472	0.55	5.18	5.44	0.41
Education	13.417	12.424	0.47	1.41	1.36	0.37
Healthcare	11.229	8.826	0.59	1.09	1.03	0.34
Law and order	21.532	17.565	0.72	2.22	1.89	0.75
<i>Local governments</i>	<i>Mean</i>	<i>Median</i>	<i>C.V.</i>	<i>Mean</i>	<i>Median</i>	<i>C.V.</i>
Total budgetary sphere	10.274	9.382	0.42	—	—	—
Public administration	30.898	27.916	0.50	—	—	—
Education	10.507	9.594	0.39	—	—	—
Healthcare	9.813	8.918	0.43	—	—	—

—, Not available.

C.V. Coefficient of variation.

Source: Calculated by the authors from the "Personnel and Client Groups" reports by about 30 regional governments.

The Main Problems in Russia's Local Government Structure

The discussion above allows us to identify three priority concerns about Russia's current structure of local government:

- *Viability of the constitutional concept of local self-government.* Following the Council of Europe principles, Russian legislation defines local self-government as the management of local affairs by local residents under their own responsibility. Moreover, the *Law on the General Principles of Organization of Local Self-Government in the Russian Federation* gives inhabitants of any settlement, urban or rural, the right to local self-government regardless of the settlement's size. Just like in some other transitional countries, these new provisions were interpreted to give any local administration the right to declare autonomy. However, after decades of citizens' passivity prevailing under the Communist regime, modern Russian society is extremely atomized. It will take some time and civic-participation experience for local residents to build social capital, including positive beliefs about their neighbors and local community and willingness to cooperate in order to improve their community. Meanwhile, local government bodies, even if somewhat more independent from the higher authority, do not seem to represent local authorities managing "a substantial share of public affairs under their own responsibility and in the interests of the local population," as stipulated in the European Charter of Local Self-Government.
- *Historical boundaries versus optimal size jurisdictions.* Ideally, structures of local governance should be based on areas that enable government services to be provided effectively and economically, thus relating area and population to local authorities' powers and responsibilities. However, until now local government has been predominantly implemented within the old administrative division of territory. The only choice is about what level of the old administrative hierarchy to make directly accountable to the community while attempting to transform the structures below into virtual deconcentrated arms of this chosen tier. It would only be by coincidence that the historical boundaries of administrative units enable optimal provision of local services.
- *Centralized pay scale.* Over the last decade, Russia has been developing its federalist principles including independent legitimacy for regional and local authorities. Among other things this independence involves discretion in the internal organization of government and personnel decisions. On the other hand, Russia inherited many traditions of centralized civil service from the previous regime. In the case of public pay, this took the form of centralized wage norms that were later transformed into a unified job classification and salary scale. The successful continuation of reforms would require accommodating these conflicting trends.

The Kozak Commission's Proposed Reforms on Territorial Administrative Division

The Kozak Commission's report issued in 2002 includes several specific recommendations on the reform of the territorial-administrative structure of local governments. The Kozak proposals essentially mandate a uniform two-tier structure of local government in the entire territory of the Russian Federation. An explicit distinction is made between rural and urban types of municipalities. At the top tier, the two types of municipalities are called municipal and urban *districts*, while at the bottom tier the names are rural and urban *settlements*, respectively. This approach would be a departure from the regional variations allowed under the *Law on the General Principles of Organization of Local Self-Government in the Russian Federation*, and appears to return to the uniform, hierarchical approach under the Soviet system.

These proposals are a clear break with the previous legislation which declared that all municipalities should have equal rights, which in turn meant that the regions could do very little to sort municipalities according to the size of their population, size of their territories, or rural versus urban status, and ultimately be able to assign revenue and expenditure responsibilities accordingly.

The Kozak proposals also regulate the minimum and maximum size of municipalities. This is again in sharp contrast to the previous legislation, which stated that no inhabited territory could be deprived of an opportunity to become a municipality. In the proposed law only those administrative units that satisfy specific criteria can become municipalities. To become a bottom-level municipality, an administrative unit should either (a) have a status of a town according to the administrative-territorial codification, or (b) represent a rural territory with a population of more than 1,000 people in one settlement or a rural territory including several settlements regardless of their population. Territories satisfying those criteria could get the status of an urban or rural settlement accordingly. There are no special minimum population requirements for municipal districts. However, there is an upper limit on the territory of a top-level municipality: the border should be reachable from the administrative center by transport within one day. Similarly, the center of a bottom-level municipality should be reachable by foot from anywhere on its territory within one working day.

It is important to note that the Kozak proposals assign expenditure responsibilities separately to each tier of local government. The proposals state that the status of an urban district could be granted in a regional law to those cities that have capacity to perform responsibilities prescribed for the top-tier municipality. All other rural and urban settlements are included in municipal districts, which provide so-called intersettlement services. Thus urban districts combine both tiers of local government and accordingly introduce taxes assigned by federal legislation to either settlements or districts.

The Kozak Commission proposes that the representative body of a municipal district, apart from being elected directly, can also be formed out of the council members of constituent settlements. In this case, constituent localities can complement the district's own-source revenue with contributions earmarked for the provision of intersettlement services (roads, transportation, healthcare, waste utilization, and so on). The contribution of each settlement would be determined on a per capita basis.

Similar to the principle maintained in the current legislation, the Kozak proposals prohibit the subordination of one local government or official to that of another tier. However, regional authorities participate in the selection of a municipal chief executive officer (CEO) if such a post is established by the municipality in addition to the elected municipal head. In this case the selection commission would have to be confirmed by the regional legislatures, which can nominate up to one-third of the commission members. The proposals also introduce an appending responsibility of subnational governments for their state (municipal) institutions and give subnational governments the discretion to determine independently the terms of remuneration for state (municipal) servants as well as employees of state (municipal) institutions.

The Kozak proposals also streamline and expedite the procedures for the removal of local officials or dismissal of local councils by the regional authorities. In addition to a normative act violating federal or regional legislation, the reasons for the dismissal of the local executive head now include any action deemed by the court to abuse human rights; threaten the unity and inseparability, national security, defensive capacity, or the legal or market system of the Russian Federation; or mismanage subventions from the federal or regional budgets. There is a time limit of three months for local councils and two months for local officials during which they can rectify their decisions in order to avoid sanctions (unless another term is specified by the court). Instead of a written warning from the regional authorities, the new law would require a second court ruling stating that the injunctions of the first court ruling have not been honored within the established time period. The Kozak proposals eliminate the previous provisions that the government of the

regional capital can be dismissed only by the federal authorities and not the respective regional authorities. In fact, the proposals suggest that federal authorities would play no role in the dismissal of local authorities even if the respective regional authorities fail to act.

Policy Reform Options

Reforming the administrative structure of the Russian Federation, for that matter that of any other country, calls for a tailored approach that recognizes the history, geography, and other specific features of the country. Nevertheless, public finance theory as well as international experience and best practices can be used to draw some conclusions and policy reform options on the reform of the structure of local governments in the Russian Federation (see appendix 3.1 for a review of the normative theory of government structure and size and appendix 3.2 for international experiences). The Kozak Commission's proposals for reform, reviewed in the previous subsection, generally move the reform agenda in the right direction, but some corrections and extensions of these proposals could be desirable. We will now discuss several key policy questions and possible options for further territorial reform.

What should the concept of "local self-government" be understood to mean? To what extent should local self-government be integrated into the "state" or government sector?

Some local governments in the Russian Federation have interpreted the concept of local self-government very broadly, and have concluded that the constitutional and legal guarantees of self-governance preclude any state (federal or regional) influence over local policies. In light of international practices, this should be considered an overly broad interpretation of the meaning of local self-government. The *European Charter of Local Self-Government* states that local governments may not be deconcentrated units of higher-level governments and that the operation of the local government should be supervised by locally elected councils. However, every country considers local government to be an integral part of the public sector, and while local governments should have substantive discretion over local policy, local governments are in some ways administratively accountable to higher-level government, whether at the regional government or national government level. In a federal system this administrative accountability should be to the regional government. By and large, the Kozak Commission's proposals adopt a good balance between the right for "local self-government" and the supervisory and monitoring activities by regional governments. It is unlikely, however, that there will be final firm boundaries between those two principles as local autonomy and regional authority will change dynamically over time. Another positive feature of the Kozak Commission's proposals is the reliance on the court system to address some of the hardest issues likely to face territorial administrative relations, such as the removal of local executives.

Should there be two levels or types of local government in Russia, or should all local governments in Russia be consolidated at the same tier of government?

The Kozak proposals mandate, as we have seen, a uniform two-tier structure of local government. It is important to recognize, however, that there is no need to have one single or uniform approach toward local government structure to be applied to all regional (oblast) governments in the Russian Federation. Economic theory and international practices suggest that different economic, geographic, and demographic conditions may be best served by different local government structures. It is true that the proposals will allow regions to sort out municipalities, when needed, according to basic factors such as rural versus urban character and population size, but the same could be achieved with a more flexible proposal that would allow the regions to adopt a model more consonant with their needs and even traditions. In fact, problems should be expected with the proposed approach. Allowing district governments to have administrative responsibility over municipalities at the settlement level in an environment

where mutual trust is lacking would likely result in instances where districts would attempt to abuse their control over the lower-level local governments in contradiction with regional law. Experiences in developing economies with multilevel governments show that each intermediate link weakens the government structure, allowing money to leak out of the system and reducing the overall accountability of the public sector.

Our comments must not be taken as a recommendation that Russia follow, for example, the American practice of allowing the states to determine fully their own local government structure. Given Russia's strong history of hierarchical governance, the absence of federal involvement may easily lead to a situation in which (some) regional governments limit the fiscal discretion of local government to their own advantage. Instead, we believe that a more practical solution for Russia would be to allow regional governments to choose from a number of approaches to local government structure. For example, in oblasts with relatively more advanced administrative capabilities, high population density, and a small geographic area, settlement governments could serve as the backbone of local government while the district-level government (if present at all) could serve more as an association of settlement governments. Then the oblast government would directly deal with each settlement's government. By contrast, in large regions with low population densities, where the delivery of local public services by small jurisdictions would result in inefficient provision of local goods and services, the districts/rayons should be the main form of local services delivery. Other models are of course possible, and consulting the regional government on these issues is likely to produce other potentially worthwhile models.

In addition, federal government policy could facilitate the creation of sound local government structures by imposing uniform national requirements for the establishment of a classification of municipalities and townships and for the formation and definition of rayons. The Kozak proposals move in this direction by regulating the maximum and minimum sizes of municipalities.

How should regions interact with different tiers of local government?

Is there a generalized need, as proposed by the Kozak Commission, for two hierarchical levels of local government in the Russian Federation? Economically there is a trade-off between the administrative ability of the regional government to deal with a larger number of municipal government units from the regional center, and the administrative costs associated with inserting an intermediate level of rayon (district) government between the region and the lowest level of local government. Economics would suggest that in very large regions, it might make administrative sense to structure the administration of local governments in a hierarchical fashion in which (for administrative purposes) municipalities report to districts, and districts report to the regional government. However, when comparing the size of the subjects of the Russian Federation with regional governments in other federal systems (for instance, states in the United States), it would be hard to fathom that *all* Russian regions could benefit substantially from such a hierarchical structure. For instance, while many U.S. states operate without an intermediate tier of local government, Russia's most populous region (Moscow City, with a population of about 9 million) would only rank as the 10th largest state in the United States.

If the hierarchical relationship between the two levels of local governments, as formulated in the Kozak Commission's proposals, were not forced onto the system, this would have two practical implications. First, it would mean that in larger regions, regional governments could opt to deal directly with a large number of municipal governments. In order to do so efficiently and effectively, regional-local fiscal relations would have to be made as transparent and formula-driven as possible (as opposed to engaging in face-to-face, negotiation-driven policies). Second, it would suggest that an important focus of regional-local fiscal reform should be to ensure improved financial management techniques at the regional and local levels that would allow regions to deal effectively with large numbers of local governments in the day-to-day management of financial operations.

Should subnational authorities have discretion over public pay?

Both the normative guidance and international experience suggest that decentralization of pay involves a number of trade-offs. Centralization of the public pay can help the central government achieve macroeconomic stabilization, preserving purchasing power of public wages and equity of pay. On the other hand, a countrywide pay scale detracts from the ability of local governments to keep more able employees or take advantage of the local labor market conditions. The pros and cons of and the international experience with subnational government civil service autonomy are discussed at greater length in appendix 3.3.

The concern about inflation control does not seem to justify maintenance of Russia's unified pay scale. During the earlier years of transition, the demands of stabilization policy might have called for a national wage policy. However, the implemented pay determination system hardly serves the goal of a national wage policy. This is mainly because of the subnational governments' discretion in supplementing the basic wage with allowances and bonuses. Moreover, the pay scale does not cover public administration employees, who tend to have pay levels above the average salary paid in the budget sphere. Currently the only possibility for the federal government to cap subnational wage bills is doing so indirectly through the determination of revenue sharing and grants among levels of government. However, with the introduction of formula-driven grant distribution at the federal and (potentially) regional levels, the significance of the pay scale for controlling the wage bill is diminished.

Currently the more important rationale for the continued existence of a unified national pay scale seems to be its role in preserving the purchasing power of public wages. This involves political concerns given the significant share of budget sphere workers in the electorate and concerns about the quality of social services such as education and healthcare. However, normative principles suggest that minimum quality standards for public services and grant conditionalities could be more effective tools than the regulation of productive inputs such as wages. Moreover, for countries of Russia's size, it is not uncommon to have price dynamics significantly varying across regions. Thus, even with the general loss of purchasing power, a uniform pay raise might not fit local circumstances.

Equity of pay does not seem to be preserved with the current pay scale either. Because of the discretionary allowances and bonuses, there is significant variation in public wages across regions and among entities of different levels of government located in one region. The only difference that the current system makes is establishing some equity of pay among employees of the same entity. However, the optimal balance between equity and performance depends on social and political values and thus can differ among localities.

While many arguments in favor of centralization of pay determination hardly apply to the Russian circumstances as indicated above, most arguments against it still hold. Thus, although subnational discretion over bonuses and allowances provides some room for managerial flexibility, performance incentives are limited as bonuses and allowances are determined on a monthly basis. Federal government pay raises destabilize local budgets because they occur within the fiscal year (most often in April and September). Federal decisions on public pay cannot accommodate variations in the local conditions of labor markets manifested in multifold wage differences. However, the major drawback of the federal involvement in subnational government pay is that it provides room for manipulation by subnational authorities.

Allegedly, the federal government's sensitivity to public wage arrears, which arose in the second half of the 1990s, led Russia's subnational governments to use perverse cash management strategies and created conditions for a soft budgetary constraint. Under this paradigm local governments appeared to prefer to allocate funds to trivial needs because urgent needs could be funded with a federal bailout later. For example, local governments may have deemed it more profitable to commit all of their resources to building a new school instead of setting aside cash for teachers' summer pay (a three-month check paid upfront). The political costs of having thousands of teachers unpaid made the federal government provide an intergovernmental loan,

which often was written off eventually. This paradigm and behavior rests on the established perception that the federal government is ultimately responsible for public pay.

The revision of the current pay system in Russia requires a careful strategy. Currently the national pay scale provides the federal government with some control over the accumulation of subnational wage arrears. Therefore if the pay scale is abolished, subnational authorities could introduce irresponsible wage raises hoping that the wage bill will eventually be covered by the federal authorities. However, subnational authorities will not be responsible for their personnel decisions as long as the federal control over subnational wage policy remains. To break this vicious circle, the federal authorities will have to come up with a credible commitment not to get involved in subnational wage crises or to make the bailout a very costly solution for subnational politicians (for example, dismissal, forced administration, harsh penalties for breaking the labor legislation, and so on).

Considering all the pros and cons, we conclude that providing autonomy for pay scales and staff management decisions at the subnational level will strengthen Russia's fiscal federalism. The Kozak Commission has basically reached the same conclusion. Civil service reform is needed to increase the productivity of the local public sector and efficiency in the delivery of public services. The reform should aim at a synchronized restructuring of careers on the basis of merit recruitment and promotion systems, with the promotion of improvements in the compensation system and an optimization of the public employment by retrenching surplus labor.

Appendix 3.1. The Theory and Practice of Local Government Structure and the Optimal Size of Local Government Jurisdictions: An Application to the Russian Federation

We noted in the main text of the chapter that the structure of local government in Russia and elsewhere is more the result of historical accident than the result of careful and sound economic design. But, of course, this statement clearly would imply that economic theory can actually provide some level of guidance in answering the question: what is the optimal size of a local government jurisdiction, and what is the resulting optimal structure of local governments?

Economic Theory and the Structure of Local Government

Economic theory has looked at these questions from a number of different angles. This appendix considers three economic concepts that are related directly to local government structure: the subsidiarity principle, club goods, and the Tiebout hypothesis.

THE SUBSIDIARITY PRINCIPLE Economists consider that the efficient provision of government services requires that government satisfy the needs and preferences of taxpayers as well as possible. This is best achieved by "subsidiarity" or the "subsidiarity principle." This principle states that the responsibility for the provision of goods and services should be placed at the lowest level of government that can efficiently deliver the good or service. In other words, subsidiarity suggests that government goods and services should be provided at the lowest level of government compatible with the "benefit area" of the service.

For instance, the benefit area for sanitation services (household garbage collection) is clearly the local community, so in accordance with the subsidiarity principle, this service should be provided (that is, delivered and paid for) at the local government level. To give another example, since the benefit area for air traffic control is the entire national territory, this service should be provided at the central (national) government level.

Leaving the supply of government services with small benefit areas up to higher-level governments may be inefficient, for the following reasons:

- The greater the distance between policymakers and the people, the less informed policymakers will be about the preferences and needs of the people.
- The central government will likely not be able to vary its mix of goods and services to match regional variations in the preferences and needs for public services.
- Central government officials will be less accountable to voters for the quality of services they provide compared to local government officials.
- Central government provision of essentially local public services will make taxpayers less willing to pay for public services, since the link between costs and benefits will be lost. (Efficiency in the provision of public services is enhanced if consumption benefits are linked to costs of provision via fees, service charges, or local taxes.)

Shifting the supply of public services with wider benefit areas to smaller units of government is likely to result in inefficient provision of services. For example, if a municipal hospital (financed by municipal taxes) were expected to provide healthcare services to the entire surrounding region, this hospital would likely provide a nonoptimal level of healthcare. Municipal residents would be unwilling to finance healthcare services for residents in the wider region; as a result of the mismatch, local residents would put pressure on their elected officials to reduce spending on healthcare services at the hospital and look for alternative means of healthcare provision.

Given the above, the subsidiarity principle argues that the structure of government should be determined in such a way that allows the lowest level of government to provide each good or service efficiently. In other words, local governments should be subdivided into smaller jurisdictions as long as the jurisdiction continues to be able to provide its goods and services efficiently. Since we cannot ignore the existence of scope economies in providing several government services through a general-purpose local government, or scale economies associated with administering a number of schools, hospitals, and the like, it might be difficult to concretely identify one single provision area that achieves this standard for all goods and services provided by a local government. Nonetheless, the subsidiarity principle is a useful tool in analyzing expenditure assignments and in determining the optimal structure of local government.

CLUB GOODS As suggested earlier, many government-provided goods, especially at the local level, are not “pure” public goods. Instead, there are government-provided services that are (largely) excludable and (largely) rival.¹⁸ Consider local provision of schools, small parks or playgrounds, libraries, and so forth. Why would the government sector provide such goods at all? Economists argue that in the case of these goods, market failure does not result from the nature of the good (after all, many goods and services that are provided by local governments are excludable and rival), but from the significant scale economies in the provision of these goods, which typically makes provision by individual consumers not feasible. A common example of a service that meets these conditions is a swimming pool; it is both excludable and (to some degree) rival, yet most households do not have a swimming pool of their own, owing to significant scale economies involved in building one. This type of issue was first explored in depth by James Buchanan, an economist.

Buchanan (1965) argued that if households would get together in clubs, these clubs could provide the relevant good to a larger number of people and lower the cost to a level that would be affordable and acceptable to everyone in the club. This, of course, is the rationale behind public pools. Buchanan thus argued that in many cases local governments simply function as

18. An *excludable good* is one whose benefits can be denied to people—if for example they are not willing to pay for it. A *rival good* is one whose use by one person detracts from the enjoyment of it by other people. Regularly, private goods are rival and excludable. Public goods, such as national defense, are nonexcludable and nonrival, and therefore the market mechanism fails to provide them optimally.

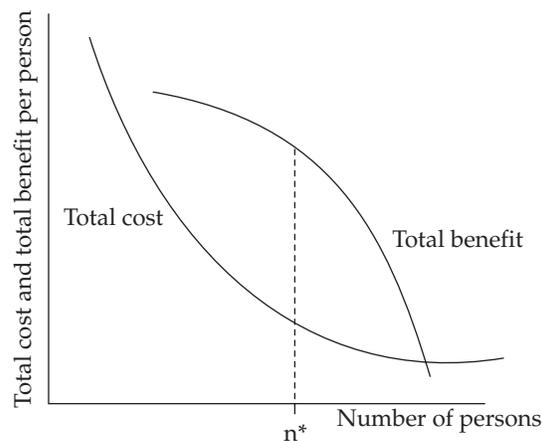
clubs that ensure that certain goods are provided and that collect user fees from households, or proxy user-fees in the form of local taxes. Many other local government goods and services are club goods, including local amenities such as parks and green spaces, markets, childcare facilities, and primary education.

Within the economic theory of clubs, what determines optimal size? In other words, what would be the optimal size for local governments to provide a certain good or service? How the optimal size of a club can be determined is illustrated in appendix figure 3.1.1, showing how the cost per person varies with the club's size (number of people in the club), as well as the relationship between benefit per person and club size. As the number of people included in the club increases, the per-person cost of providing the good or service declines as a result of scale economies. However, as the club's size increases, the benefit to each club member decreases as more people share the facility (depending on how rival the good is and the "congestion costs" incurred), and as some people now may have more limited access to the facility as the coverage area grows (and transportation costs increase). The optimal size of the club occurs where the difference between costs and benefits is maximized at the number of people equal to n^* .¹⁹ The theory of clubs further suggests that there is a point beyond which the total cost of a club exceeds the total benefit; establishing clubs (in our case, local governments) beyond that range would thus *reduce* the welfare of its residents.

THE TIEBOUT HYPOTHESIS One problem of Buchanan's theory of clubs is that it is largely silent on how clubs are formed across a given geographic space and how they are governed. If people with diverging demands for club goods located near one another, it would make it unlikely that clubs (local governments) would actually form, as there would be no consensus on what level of club goods should be provided. In addition, Buchanan's theory is silent on how clubs (local governments) would be governed, so that there is no assurance that local governments are managed in a way that promotes efficient provision of local government services.

Charles Tiebout built on the notion that local governments are akin to clubs formed in geographical space, and explored the efficiency properties of a system of local governments. Tiebout's hypothesis (1956) argues that the provision of local government services in a metropolitan region with a large number of local governments will result in allocative (consumption) as well as technical (production) efficiency under certain assumptions.

Figure 3.1.1 The Optimal Size of Economic Clubs



Source: Authors.

19. In terms of microeconomic optimization, this occurs where the marginal benefit of changing the club's size equals the marginal cost of doing so.

The crux of the Tiebout hypothesis is that households will try to maximize their level of satisfaction by locating in the local government jurisdiction that provides the level and quality of local government services that they desire, and by locating in the local government jurisdiction that provides these local government services in an efficient manner. Local governments provide local public goods and cover the costs through (proxy) user fees.

A first implication of the Tiebout hypothesis is that residents will sort themselves in clubs with homogenous preferences for local public goods, since all residents within a jurisdiction receive the same level of public service and pay the same price. The Tiebout hypothesis suggests that households “vote with their feet.” For instance, if a local government offered an unattractive basket of public goods and services, this household would identify a jurisdiction that provides exactly the package of local government services the household desires and would relocate.

A second implication of the Tiebout hypothesis is that competition between jurisdictions prevents local governments from being inefficient or from engaging in redistributive policies (not supported by taxpayers). For instance, if a local government were excessively inefficient in the provision of local public services, households and firms over time would move to jurisdictions where their local tax resources are not wasted. Similarly, if a local government engaged in redistribution, wealthier households would “vote with their feet” and leave the jurisdiction, thereby reducing the local government’s tax base and its ability to engage in redistributive policies.

One of the policy implications of the Tiebout hypothesis is that regions with a greater number of local governments should achieve a greater level of allocative efficiency and be more efficient in providing local public services—because of the increased ability of local governments to match the local needs and because of the increased competition by local jurisdictions. A precondition, however, for such efficiencies is that households be able to assess the relative (in)efficiency of their local government when making location decisions. As such, the Tiebout hypothesis implicitly requires transparency in the operation and administration of local governments.

Empirical Evidence

For policy purposes, concepts and theories are only of limited value if they are not able to link back to the reality in which government policies have to be formed. What concrete guidance does the theoretical framework provide for the policy discussion on the optimal size of jurisdictions and optimal local government structure?

In the foregoing discussion, scale economies have been assigned an important role in the determination of an optimal structure of local government. In fact, a postwar series of local government boundary reforms across much of Europe was driven by the idea that government efficiency would be increased by the creation of large authorities. The efficiency was understood in a sense that the same level of public good is produced with lower costs. However, from an empirical point of view it has proven difficult to define the costs of public services, as it is very difficult to account for qualitative aspects and externalities in public service provision. Nevertheless, many countries undertook studies to determine the optimal size for local authorities by measuring only direct costs and assuming that the level of service provision is invariant. Most of these studies failed, however, to produce a clear-cut answer on the optimal size of jurisdiction (Martins 1995). The optimal size seems to differ by type of service and thus requires assigning weights to different functions of a multipurpose government. The best that such studies could do was to help determine local authorities’ minimum viable size threshold for a given basket of services. Depending on the functions of local government in a particular country, this threshold varies from 2,000 inhabitants in Bavaria to 5,000–6,000 inhabitants in Denmark, 6,000–7,000 in the Netherlands, and 8,000 and 15,000 inhabitants of rural and urban areas, respectively, in the Saarland (Germany).

The empirical literature on the provision of local government goods is too extensive to be adequately summarized here. However, there seems to be a consensus among economists on a number of issues. The general conclusions of the literature appear to be the following:

- Scale economies exist in the provision of local government goods and services, but the level of scale economies differs greatly for different kinds of locally provided goods, and the level of scale economies differs greatly in different settings. Typically it is possible to determine a minimum size threshold for a local government that is needed for a government unit to be able to produce a certain good or service efficiently.
- When several local jurisdictions are available, households display a tendency to sort themselves with regard to income and/or the demand for local government services. However, even with large numbers of competing jurisdictions available, large within-jurisdiction variations in income and preferences continue to exist.
- As the Tiebout hypothesis suggests, it appears that local government services (such as local public schools) feel some competitive pressure to provide services in an efficient manner. However, elections and “households voting with their feet” (by moving out of inefficient jurisdictions) are slow and imperfect mechanisms, allowing considerable inefficiencies to occur at the local level, even when multiple alternative jurisdictions are available.

In summary, the main conclusion regarding the optimal size of jurisdictions and the optimal structure of local government should be that a single approach toward government structure will not fit all situations. While economic theoretical considerations can provide some guidance in framing the overall discussion, the optimal structure of local governments can vary from region to region and country to country, and will depend on the kinds of goods and services provided by local governments, the cost structure of the local government–provided goods and services, the political and institutional environment, and the demographic and geographical conditions.

Scale Economies and Alternative Modes of Producing Local Government Goods

The existence of scale economies is often used as an argument (or a pretense) to prevent devolution of an expenditure responsibility to a lower level of government. Since subnational government units at any local level typically vary in size, it is not uncommon for opponents of decentralization reform to argue that especially the smaller local governments are not able to take on the burden of government services that display a certain degree of scale economies. Such concerns have unnecessarily hindered the implementation of decentralization reform since there are several ways in which jurisdictions (particularly smaller ones) may circumvent the negative consequences of scale economies.

The crux of the argument here is that just because local governments have been assigned the responsibility to fund and *provide* a certain government service, this does not necessarily mean that they have to *produce* this good or service themselves. There are three ways in which a local government jurisdiction can provide a good or service without producing it on its own and still retain the benefits of local policy discretion and accountability:

- *Contracting with the private sector.* A local government may choose to contract with a private provider to produce a certain local government service, such as garbage collection or the operation of the local market. By doing so, the benefits from fiscal decentralization (including a better match between service provision and local preferences, as well as increased local accountability) are retained, while also being able to capture the scale economies and experience brought to bear by large, specialized private service providers.
- *Contracting with other local governments.* Another possibility to avoid the trappings of scale economies is for two smaller jurisdictions to collaborate on the production of a certain public good. For instance, one local jurisdiction that is too small to provide secondary schooling in a cost-effective manner on its own might contract with a neighboring local government and send its students to the neighboring district. Alternatively, the two districts might establish a covenant to jointly run a secondary school for the education of children from both jurisdictions.

- *Forming special service districts.* In certain cases, special service districts could be formed for the provision of certain local public goods. This is particularly the case in situations in which several local governments are inextricably linked by the provision of a certain local government good or service, such as a water district that covers several local government jurisdictions. Again, special districts allow a system of intergovernmental relations to resolve problems with scale economies without altogether abandoning the benefits of a decentralized system of local government.

Democratic Governance and Local Government Structure

One influence on the structure of local government that has not been adequately addressed so far is the political desire to have participatory and democratic governance processes. Policy experts need to take into account this political dimension of local government structure, which is often overlooked by economists. As such, historical precedent and institutions cannot be ignored altogether, but neither can the perception that smaller government jurisdictions are inherently closer to the people and thus more democratic. For example, based on data from Denmark (both at the municipality and individual levels), Mouritzen (1989) examines the impact of size on attitudes toward local government. He finds that citizens in small jurisdictions hold more favorable attitudes toward participation and democracy, and he further finds that small local government units are more homogeneous and in fact more efficient in the provision of services.

In the end, the economic theories that frame the question of optimal government structure do not provide a conclusive answer on how many levels of government should exist in a country, or at which level (for instance, district, town, or village) local governments should be established. The conceptual framework outlined in this appendix, including location theory, subsidiarity, the theory of clubs, and the Tiebout hypothesis, suggests that there are trade-offs in determining the optimal structure of local governments, between positive and negative features of adding an additional level of government. On the one hand, a local government structure that predominantly relies on smaller jurisdictions may fail to capture economies of scale in the provision of public goods and services; on the other hand, relying on larger jurisdictions will reduce the level of allocative efficiency (matching households' preferences with the level of service provided) and competitiveness between local governments. Relying on multiple levels of local government, several types of local governments, or even special-purpose local government also comes at a cost, often causing duplicative administrative effort (thus failing to capture scope economies) and increasing the information cost for households; under this scenario households would have to spend more time being informed in order to hold governments accountable.

The arguments in favor of and against adding an additional level of government in a government structure (or further fragmenting the lowest level of government) are contained in appendix table 3.1.1.

As we concluded above, while the empirical literature is generally supportive of the conceptual framework and its policy implications, there is little empirical literature that provides concrete guidance in handling the trade-offs. In conclusion, the determination of what constitutes an optimal structure of local government will have to rely heavily on the existing institutional framework and a review of international practices and experiences.

Appendix 3.2. International Experiences in Structuring the Local Government Sector

Many of the questions that relate to the structure of local government in Russia are a matter of legislation or regulation. For instance, how should local governments be formed—bottom-up (local voting and local charters) or top-down? If top-down, should it be the federal or the oblast

Appendix Table 3.1.1 *Potential Positive and Negative Effects of Adding an Additional Level of Local Government*

<i>Potential pros</i>	<i>Potential Cons</i>
<ul style="list-style-type: none"> • Achieves more homogeneous jurisdictions at the lower level • Better match in provision areas for goods with small benefit areas: improves allocative efficiency • More competition between local governments (see Tiebout 1956), which increases administrative efficiency • Government gets closer to the people, which leads to increased political participation and a better informed consumer of local public goods 	<ul style="list-style-type: none"> • May decrease scope economies in administering public goods, leading to technical inefficiency • Excessive numbers of government levels may increase information costs and reduce participation, transparency, and accountability • Fragmentation may ignore scale economies in provision and administration, leading to technical inefficiency

Source: Authors.

level that controls local governments? Should there be a federal law strictly governing the formation of local governments, or should this be left up to the states? What expenditure responsibilities are assigned to the local level, and what discretions do local governments have? All of these questions address the question of how the structure of local government will truly be formed. In addressing these policy questions, a lot can be learned from the practices and experiences of other countries.

International experience presents two general approaches to local government: community-based local government and utilitarian administration.

Community-based local government is understood as one that is made up of citizens of the community—"the people who belong to its area and have the right to participate in governing themselves as a collectivity" (Norton 1994, p. 17). This approach is common for Council of Europe countries and is especially popular in Denmark and the Nordic countries. In those countries, local governments also serve the purpose of the "social state," by intervening to correct malfunctions in the market, giving assistance to the disadvantaged, and observing the rights of labor. Local self-government or home-rule is also the norm in many U.S. states for municipal governments (Katz 1999).

By contrast, "utilitarian administration" refers to the body that in fact exercises power over the citizens, educates them, and creates opportunities for self-realization. This approach, prevailing in Anglophone countries, complements a laissez-faire state, which is justified in terms of individual choice and economic efficiency.

In Central and Eastern European countries, although there were some systems of local government before the transition, all power structures had been mostly authoritarian and provided little opportunities for self-government. Therefore the process of democratic institution-building required the creation of a new system of local government. In this study we survey six countries in transition: Bulgaria, the Czech Republic, Hungary, Lithuania, Poland, and Slovenia. All transitional countries surveyed here adopted the Council of Europe approach to local governance and signed the *European Charter of Local Self-Government*.

The Status of Local Government

In the Council of Europe, countries' written constitutions guarantee a system of local governance and give it a secure position within the national system. In the United Kingdom, although the rights of communities to self-government has received general recognition, the "unwritten constitution" affords no protection to them. This became clear under the Margaret Thatcher administration, which in various ways lowered the level of discretion available to local authorities and prevented them from increasing the quantity and quality of their services.

Similarly, federal constitutions of the United States and countries in the British Commonwealth (including Australia and Canada) make no provisions for local government. However, constitutions of constituent states typically do recognize local authorities and protect them from arbitrary dismissal. In the continental Europe, constitutions of member states of federations and regional governments often detail competences of local governments, as in Switzerland, Germany, Italy, and Spain. In Japan, the 1947 Constitution recognizes local government as a part of the new democratic administrative system based on the rule of law.

Structure of Local Government

A general tendency observed internationally is that countries with a utilitarian approach to government favor large-scale single-tier local authorities. At the same time, countries with community-based local government tend to have complex multi-tier systems with a large number of small, basic authorities (Norton 1994). Also, three out of the six transitional countries have intermediate local authorities in place and Slovenia is preparing to create the second tier of local self-government. In the Czech Republic and Hungary, new local government legislation has been exploited by thousands of small villages to claim the status of an autonomous municipality. Previously, the reaction to forced consolidations and deprivation of services and development under Communist regional planning policies had resulted in high fragmentation of local authorities by the late 1990s. In the Nordic countries, where consolidations created large basic authorities (in Sweden, for instance), submunicipal authorities with advisory or executive powers were established later.

In a number of federal countries, structures are mixed in the sense that the local government structure and number of local government tiers vary among different regions of the country (for instance, different states in the United States have different local government structures). Nevertheless, in general the structure of local government seems to be related to the concept and role of local government in a particular country.

As we saw above, both single-tier and multi-tier structures bring about different benefits. To recapitulate, the benefits of single-tier all-purpose government are as follows:

- Economies of scope (coordinated provision of different services);
- A single counterpart to the central government (one strong voice is more effective than many weak players); and
- Horizontal accountability (there is no confusion of what authority to hold responsible for the quality of services).

The benefits of having multiple tiers are as follows:

- Decisionmaking can be kept close to the people if intermediate authorities assume responsibility for services that are inefficient when provided at a small scale.
- Intermediate authorities can afford large-scale investments (for instance, in specialized equipment).
- The scale and level of specialization of government agencies can be increased.
- Losses to horizontal accountability that would result from consolidation, joint authorities, or centralized provision are avoided.²⁰
- There is more contestability of power, owing to the larger number of independent authorities.

International practice offers several arrangements for the intermediate local authority (Norton 1994, p. 42):

20. For instance, in Bavaria for municipalities within the area of jurisdiction of an intermediate authority (*Kreis*) the minimum population is fixed at 1,000 inhabitants, while for unitary municipalities this figure is 2,000 (Martins 1995).

- A state-appointed official responsible for state functions over a large area without an elected council of its own (examples of this include Bulgaria and Lithuania);
- A state-appointed official for each intermediate local jurisdiction who acts as executive and presides over an elected council; and
- A prefect who is responsible for state functions in the intermediate local jurisdiction in parallel with an autonomous assembly that provides major services and elects its own chairman and executive.

Unlike the former Soviet system, in established democracies multi-tier systems of local government are generally not hierarchical—that is, there is no political control of one tier over another. And in some transitional countries (such as Hungary, Poland, and the Czech Republic) tiers of local government are nonhierarchical.

While in many of the developed countries it is now a principle that all tiers of local government enjoy a fairly high degree of autonomy, relationships among different levels of government are still quite close in many unitary countries in continental Europe (Norton 1994, p. 48). This has been brought on by persuasion, incentives, and sometimes legal requirements. The *Local Government Law* in Italy, for example, repeatedly requires interlevel consultation. In France, communes have been obliged—often by fiscal pressure from the State—to join together in single- or multi-purpose syndicates sometimes vested with taxation powers. In Sweden, local government legislation requires sharing responsibility among authorities for regional and local transportation usually carried out through jointly owned companies. In Finland, joint municipal boards are established with responsibilities for hospitals, care of the disabled, and regional planning. In the Nordic countries, local authorities readily cooperate with one another and with other institutions to provide services for their residents (Hughes and others 1998). Finally, in each of the 58 state administrative units in Slovenia, municipal councils appoint their representatives to the advisory committee that ensures coordination between municipal bodies and state administrations.

The *European Charter of Local Self-Government* states that any administrative supervision of local authorities may only be exercised in cases determined by the statute and aim only at ensuring compliance with the law or with regard to expediency in respect to delegated tasks. Administrative supervision has been reduced to a simple supervision of legality in Spain, France, Portugal, and Sweden. In principle this is also the case in Denmark, Finland, Norway, and Poland, but exceptions are tolerated in particular sectors or under special laws (Denmark, Finland) and in matters concerning “delegated powers,” as in Poland (Council of Europe 1994). When supervision of expediency is envisaged, considerations of “economy” (Poland, Austria), “efficiency” or “utility” (Austria), or “strictly necessary situations” (Netherlands) may intervene.

Larger federal countries often have a need to be more flexible in their approach to local government structure, as a single approach may not meet the needs of different regions within the country. In some federal countries (including the United States and Canada), it is completely left up to each state government to define a local government structure. Thus it is left to the states to ensure that local government services are provided across their state’s territory, and to accommodate whatever degree of local self-rule that is expected by local communities. Owing to similarities in the size and diversity of the countries, the general framework of local government structures in the United States might hold important lessons for the Russian Federation.

Most states in the United States rely on one level of local government consisting of several tiers or types of local government units. In virtually all states, the first tier of local government is a system of county (district) governments, which can be taken to be equivalent to rayon governments in Russia.²¹ The number, size, and boundaries are typically defined by state legislation,

21. In some states, counties are known by different names; for instance, counties are known as “parishes” in Louisiana. Several small states in the northeastern United States do not have county governments at all. Hawaii and Alaska have local government structures that in many ways are unlike the local government structures found in the 48 continental United States.

but counties are governed by locally elected boards. The country structure ensures that local government services are provided across the state's entire territory, including areas that are not covered by municipal or other local governments. The United States comprises 50 states (average population: 5.5 million) and has 3,043 counties (average population: approximately 90,000).

In addition to the county government type, many U.S. states rely on a lower tier of local government, namely municipal (city) or township (town or village) governments, to provide local government services. Every state has its own legislation on how municipalities and townships are formed and how municipal boundaries can be changed. Typically, municipalities and townships are established through the local electoral process, allowing the formation of self-governing municipalities and townships depending on a number of characteristics, such as minimum population, minimum geographical area, and minimum population density. Subsequent to their establishment, municipalities and townships are governed by locally elected officials, on the basis of a ("corporate") local charter, which is drafted in accordance with the state's local government legislation.

The United States has a total of 19,372 municipal governments (median population: approximately 75,000) and 16,629 township governments (median population: approximately 25,000). In addition to these general-purpose local governments, the local government structures in most states rely heavily on special-purpose local governments, including a total of 13,726 school districts and 34,683 other single-purpose districts such as fire protection districts, water districts, and housing districts. These school districts and special-purpose districts are generally run by separately elected boards (such as school boards), and their boundaries may or may not coincide with county or municipal boundaries. To complicate matters further, in some cases county and municipal governments combine to form a single entity, such as Miami-Dade County, the city-county entity that resulted from the merger of the City of Miami and Dade County in Florida.

It is important to recognize that municipal and township (or special-purpose) governments in the United States are part of one level of local government and not hierarchically organized; neither municipal and township governments nor school districts are in any way accountable to the county government. Instead, they all deal directly with the state government; state law in each state defines the rights and responsibilities and the revenue sources of each type of local government. These assignments typically require county governments to provide all local government services outside the boundaries of municipalities and townships, and assign municipal and townships responsibility for most or all local expenditures within their boundaries.

The fact that local government is not organized in a hierarchical fashion does not mean that county and municipal governments do not interact, however: while local governments pursue individualistic, competitive objectives, they form a bargaining relationship with each other and with the private sector. For instance, in many cases municipalities bargain with their county government to provide certain services in their municipality while retaining others; for instance, a municipality may exert local control over land-use zoning and other municipal functions but rely on the county for police services or fire protection (at a negotiated price, of course). In other cases, county and municipal services simply coexist, such as a county's Sheriff's Office with a municipal police department. The county is the most frequent provider of services to other levels. For example, the municipality of Lakewood is contracting out almost all of its services to Los Angeles County. In many other cases, municipalities also rely on the county to collect their local taxes rather than setting up their own local tax administration. For instance, through a contract, the City of Atlanta relies on Fulton County to collect its municipal property taxes; in return, the city pays a fee of 1 percent of property tax collections to the county.

While the approach taken in the United States allows for local government structures to adjust to each state's needs, there are two main disadvantages of allowing states full control over state-local relations. First, leaving local government structure completely up to the state may allow state governments to "abuse" their control over the local government level by

imposing unfunded mandates and not providing adequate local autonomy. For instance, local governments in Canada have been complaining of this in recent years and have called for federal intervention to protect local governments from the abuses of state governments. A second important disadvantage of the U.S. approach is that the system is excessively complex and lacks any degree of uniformity as no two states are alike. This has made it extremely hard to systematically collect and report statistics on local government expenditures and local government performance.

One “redeeming” characteristic of the U.S. federal approach in response to this lack of uniformity has been the imposition of a uniform data standard for reporting local government information. A key feature of this system has been the Federal Information Processing Standard (FIPS) coding system, which is a systematic coding scheme for all government units. For instance, each state government is given a two-digit numerical code. Each county government within each state is given a three-digit numerical code, thereby allowing each county within the United States to be identified by its five-digit (state plus county) code. Furthermore, each municipal, township, or special-purpose government within each state is also assigned a more specific, unique code, thus allowing each government unit in the United States to be easily identified for statistical purposes.

The Size of Jurisdictions

Several trends can be observed in different parts of the world with regard to the size of local governments. During the early 1950s and up until the mid-1970s, many Northern European countries undertook a consolidation of local jurisdictions in order to establish a minimal size for basic authorities: 8–10 thousand inhabitants (Norton 1994, p. 38). This reorganization was prompted by the perception that new public-services technologies required higher levels of public investment. In addition, setting a meaningful size for local authorities was seen as a condition of decentralizing the power from the center. However, after the consolidation took place, advisory and executive authorities were created at the submunicipal level to bring public administration closer to the citizenry.

On the contrary, in Southern Europe, local opposition to consolidation, sometimes supported by court rulings, preserved small authorities. In those countries, (small) local governments had to achieve technical efficiency by means of intermunicipal cooperation, upward delegation of functions, and contracting-out of services. It is suggested that Southern European countries (France, Greece, Italy, Spain) are less prone to restructuring because they have a fused hierarchical system of central-local relations, in which the delivery of local services is often ensured by central government field offices (Martins 1995). Out of the six transitional countries, Lithuania undertook reforms that abolished the lowest level of the Soviet hierarchy and established municipalities at the rayon level in 1995.

As of the late 1980s, many developed countries vary widely in the sizes of their basic authorities (appendix table 3.2.1). France has the smallest average size of basic authority (1,560 persons) and the United Kingdom has the largest average size (122,000). Germany and the United States stand between Southern and Northern Europe in the size of their local authority (7,000–9,000). In those European countries that have large basic authorities (for example, Sweden, Denmark, and the Netherlands) submunicipal area councils provide advice on local needs and planning issues and in some cases undertake minor responsibilities. Similarly, in the United States and the United Kingdom, local administrative offices were created in a number of cities to establish closer consultative arrangements with local residents. Transitional countries also show significant variation in the size of the average population of the lowest tier of local government. The smallest sizes are found in the Czech Republic (1,659 persons) and Hungary (3,241 persons). At the other extreme, the average size of basic authority in Bulgaria is 31,600 persons, while in Lithuania it is 66,300.

Appendix Table 3.2.1 World Systems of Local Government

<i>Southern Europe</i>						
	<i>France</i>	<i>Switzerland</i>	<i>Austria</i>	<i>Spain</i>	<i>Italy</i>	<i>Greece</i>
<i>Constitutional status</i>	<i>National constitutional</i>	<i>State constitutional</i>	<i>National constitutional</i>	<i>National constitutional</i>	<i>National constitutional</i>	<i>National constitutional</i>
Number of subnational tiers	3	2	2	3	3	2
Average size of basic authority (thousand persons)	1.560	2.122	3.421	4.997	7.019	11.225
Average size of intermediate authority	556	n.a.	n.a.	809	611	201.1
Average size of regional authority	2,521	248	896	2,380	2,832	789.1
Powers	General competence and statute	General competence	Statute	General competence and statute	General competence and statute	General competence
Control of legality by	State and courts	Courts	State	State and courts	Courts	State
Local executive authority of the community	Council	Board of managers	Board of the municipality	Mixed	Council	Mayor or president
<i>Northern Europe</i>						
	<i>Germany</i>	<i>Norway</i>	<i>Finland</i>	<i>Denmark</i>	<i>Netherlands</i>	<i>Sweden</i>
<i>Constitutional status</i>	<i>National constitutional</i>	<i>Creature of parliament</i>	<i>National constitutional</i>	<i>National constitutional</i>	<i>National constitutional</i>	<i>National constitutional</i>
Number of subnational tiers	3	2	1	2	2	2
Average size of basic authority (thousand persons)	8.845	10.000	11.206	18.811	27.230	30.249
Average size of intermediate authority	170	229	n.a.	365	1,241	358
Average size of regional authority	4,905	n.a.	n.a.	n.a.	n.a.	n.a.
Powers	General competence	General competence and statute	General competence	General competence and statute	General competence	General competence and statute
Control of legality by	State and courts	State and courts	State	State and courts	State	State and courts
Local executive authority	Mixed	Council	Council	Mixed	Appointed mayor	Mixed

Appendix Table 3.2.1. (cont.) World Systems of Local Government

<i>Central and Eastern Europe</i>						
	<i>Czech Republic</i>	<i>Hungary</i>	<i>Slovenia</i>	<i>Poland</i>	<i>Bulgaria</i>	<i>Lithuania</i>
<i>Constitutional status</i>	<i>National constitutional</i>					
Number of subnational tiers	3	2	1	3	2	2
Average size of basic authority (thousand persons)	1.659	3.241	10.344	15.561	31.614	66.3
Average size of intermediate authority	82	538.4	n.a.	82.8	n.a.	n.a.
Average size of regional authority	738	n.a.	n.a.	2414	936	371.2
Powers	General competence and statute	Statute	General competence and statute			
Control of legality by	State and courts	State and courts	Courts	State and courts	State and courts	State
Local executive authority	Municipal board	General assembly	Mayor	Administrative board	Mayor	Executive board
<i>Anglophone Countries and Japan</i>						
	<i>Canada</i>	<i>United States</i>	<i>Australia</i>	<i>Japan</i>	<i>New Zealand</i>	<i>United Kingdom</i>
<i>Constitutional status</i>	<i>State constitutional</i>	<i>State constitutional</i>	<i>State constitutional</i>	<i>National constitutional</i>	<i>Creature of parliament</i>	<i>Creature of parliament</i>
Number subnational tiers	Mixed	Mixed	Mixed (mostly single-tier)	2	2	Mixed
Average size of basic authority (thousand persons)	5.594	7.608	19.114	37.000	46.729	121.727
Average size of intermediate authority	n.a.	90.000	n.a.	37.200	234	404
Average size of regional authority	2,533	5,008	2,495	n.a.	n.a.	5,100
Powers	Limited by statute	Limited by statute	Limited by statute	General competence	Limited by statute	Limited by statute
Control of legality by	Courts	Courts	State	State and courts	Courts	Courts
Local executive authority	Mixed	Mixed	Mixed	Mayor or governor with board	Chief executive	Council

n.a. Not applicable.

Source: Norton(1994); Horváth (2000); Wetzel (2002).

These averages hide tremendous variations in the scale of basic authorities within each country. Thus, within all countries with the size of basic authority below 10,000, there is a large gap between a few dozen large municipalities (more than 100,000 inhabitants) and a vast majority of very small entities (fewer than 5,000 inhabitants).

Boundaries

Practices and legal procedures for changing local boundaries vary tremendously across countries. Legislative provisions stipulating prior consultation exist for example in German, Danish, Finnish, French, Greek, and Polish legislation. These consultations usually take place in practice, and resistance to attempted consolidation, in particular, can be expected. The consultations often produce no more than a simple opinion, which can be overridden by the competent authority. The consultation may concern the municipality itself (Spain, Poland, Portugal), the citizens involved (Denmark, Italy, France), or simply an expert committee to the prefect (Greece) or a special council to the Minister of Interior, comprising members nominated by representatives of the two municipal levels as in Denmark (Council of Europe 1994).

All countries that have carried out major structural reforms (the Netherlands, Scandinavian countries, and the like) have done so on a compulsory basis, usually after the encouragement of voluntary consolidation had failed (Martins 1995). In the Netherlands, the Parliament has the power to restructure local government, and it has often done so (Hughes and others 1998). In Switzerland, the communes have an unwritten constitutional right to merge with other communes or to remain independent—a freedom that cannot be withdrawn by cantons. In New Zealand, the Local Government Commission was established to preside over the reforms of local authorities' internal structures as well as their boundaries. In Slovenia, the National Assembly (Parliament) legalizes the establishment of new administrative units along with the demarcation of their territories following a local referendum in the affected area. And in Bulgaria, any decisions on the creation, transformation, or dissolution of municipalities, including any changes in boundaries, are subject to local referendum.

Asymmetry

The design of central-regional fiscal relations has recently involved more discussion than in the past about the pros and cons of symmetric versus asymmetric policies: should a federal or central government treat all state-level governments the same? Two related questions arise at the regional-local level. First, should each region follow the same approach to local government structures? Second, should regional governments treat all local governments the same, or would there be reasons to treat different local governments in a different fashion? This issue becomes especially relevant in light of the great differences in administrative capabilities between strong urban local governments and weaker rural local governments.

With regard to the first question, many federal structures have opted to allow regional governments varying degrees of discretion in determining local government structure. As already mentioned, in the United States and Canada, local governments are creations of the state government. In Germany, *Länder* (states) have less discretion but still a variety of local government structures exist: some regional governments have a local structure with municipalities of a small average size (for example, Mecklenburg-West Pomerania), while others have a local government structure with relatively large local governments on average (for example, North-Rhine Westphalia).

With regard to the second question, whether regional governments should treat all local governments the same, international practices again reveal a variety of approaches. Two policy choices can be defined in this regard. In one case, all local governments would have exactly the same expenditure responsibilities. This of course would require a relatively high degree of consistency

in the size and administrative capabilities of local governments, a case likely reserved for small, homogenous unitary countries. In the other case, local governments with different characteristics would be classified according to a given scheme and assigned different expenditure responsibilities and revenue-raising authority based on their size and administrative capabilities. The latter is the common approach among international practices. For instance in Spain, legal provisions assign different degrees of responsibility to local governments based on their size.

Appendix 3.3. Is There a Need for Civil Service Reform at the Subnational Level in the Russian Federation?

Local government decisions concerning employment and salary structure involve one of the most complex issues in public-sector management. No fiscal decentralization is complete without some degree of administrative decentralization, and the biggest step in this process is granting local governments authority to fire and hire their employees and to set salary scales independently of the central authorities. Because payroll is easily the largest item in local budgets, hiring and salary decisions may have strong impacts on the fiscal viability of local governments.

International experience presents a wide range of approaches to civil service management. In some countries, all public personnel decisions are made by a few power bodies in the central government. For example, in China, the Committee for Public Service and the Personnel Bureau determines staffing levels for all agencies. Prior to 1983, Singapore's Public Service Commission would appoint, confirm, and place the permanent or "pensionable" position, and promote, transfer, dismiss, and exercise disciplinary control over a majority of public officers. In the Republic of Korea, the Ministry of Government Administration is responsible for all personnel matters related to public servants, including determination of salary levels and the establishment of various categories of occupations and grades for each government authority.

In Western democracies, hiring and promotion activity tends to be decentralized at least to the ministry level. However, in many Organisation of Economic Co-operation and Development (OECD) countries there are unified qualification criteria for all civil servants of the central government, sometimes extending to the subnational levels. More variation is observed in the determination of pay and promotion. Federal countries typically have independent pay systems at each level of government. This also includes federations among developing countries such as India. On the other hand, in some unitary countries, local governments are treated the same way as central government departments with regard to pay determination (for example, France).

Recently, however, in most OECD countries there has been a trend to delegate pay determination to the department or workplace level. As a result of this decentralization trend, local governments have been obtaining salary-setting or pay-determination authority even in unitary countries.

Over the last decade, Russia has been developing its federalist principles. Among those principles there is one on the independent legitimacy of regional and local authorities. This independence involves, among other things, discretion in the internal organization of government and personnel decisions. On the flip side of the coin, Russia has inherited many traditions of the centralized civil service of the previous regime. In the case of public pay, the country inherited centralized wage norms and transformed them into a unified job classification and salary scale. The successful continuation of reforms requires Russia to accommodate these conflicting trends.

In this appendix we review the pros and cons and the international experience with local government civil service regulation with the objective of extracting useful lessons for subnational administrative decentralization reform in the Russian Federation.

Normative Considerations: The Pros and Cons of Central Control

First, here are some of the arguments that have been used in favor of central government intervention and control of salary setting at the subnational level.

PRO: MACROECONOMIC STABILIZATION AND STRUCTURAL ADJUSTMENT Macroeconomic considerations may require containment of public expenditures and adjustment of wage differentials between the public and private sectors. Because of its high share in public expenditures, the public wage bill can be used to control inflation through its effect on the money supply. Thus, for example, the austerity measures introduced by the United Kingdom after the 1976 IMF loan included cash limits on expenditures in various parts of the public sector (Silvestre and Eyraud 1995). This was implemented either directly as in the case of the national civil service or indirectly, by limiting government grants to local authorities or the universities and imposing limits on debts that state-owned companies could incur. However, the stricter application of the pay policy in the public sector led to distortions in the traditional pay structure and increased militancy among public sector workers and their trade unions. Thus, although public sector pay fell during the restrictions of 1976–79, it moved forward again during the catching-up period between 1979 and 1982. This catching-up was due to the increases awarded by the Labor government commission, established in 1979 in order to address the relative decline in pay in the health service and education sectors by applying the norms of wage comparability with the private sector.

Other arguments have been developed in the context of structural adjustment policy, defined as a process of achieving rapid structural change in an economy, so as to secure stable growth of national income over the long term (Colclough 1997). To the extent that public-private wage differentials provide incentives for the domestic production of traded goods, public sector pay can be a tool of structural adjustment. This is because, in response to prolonged deficits in the balance of payments, the major aim of structural adjustment programs has been to increase the production of tradable goods and services, relative to those of nontradable goods.

Another argument for public pay control states that the slow progress made by many developing countries has been caused mainly by the excessive economic intervention of their own governments. This point of view argues that the costs of this intervention have been large, and that the price distortions to which it has led have resulted in the market's being hindered in promoting efficient resource allocation. It is implied that for countries with unsustainable imbalances on external account, a successful adjustment would be unlikely in the absence of some reduction in the level of government spending relative to the gross national product (GNP). The wage and salary bill—accounting, as it does, for a high proportion of total public spending—represents one of the main ways in which such expenditure reductions may be secured.

PRO: PRESERVING PURCHASING POWER UNDER HIGH INFLATION In the environment of high inflation, demands for public sector raises mostly concern declining real wages. Across-the-board wage increases are demanded to compensate for the general loss of purchasing power of public wages. Case studies of developing countries show that the erosion of public wages can have disastrous effects on the productivity of civil servants (Colclough 1997). However, being the largest employer, the state can aggravate inflation by accelerating the nominal wage race. Moreover, across-the-board raises often lead to a compression of public wages, thus hampering productivity with weak incentives.

In OECD countries, one key element of the changing situation with public sector pay is the low level of inflation recorded in recent times. This new factor has made it possible to break with the old collective wage-bargaining tradition and has given the various negotiating partners the opportunity to reason more in terms of strategy. More and more, the point at issue has been the creation of a link between pay and performance or negotiations over conditions rather than wage-raise demands for greater purchasing power. The productivity requirements gradually being extended to public services are also helping to give rise to new staff management techniques, new pay schemes, and new negotiation agendas.

In the United Kingdom, from 1981 onwards, the calculation of the cash limits for government agencies included the estimated rise in wage costs. However, on September 15, 1981, the Chancellor of the Exchequer announced that “the pay bill does not mean that all increases in

the public service will be 4 percent. Some will be higher, some will be lower. There is no entitlement to an automatic increase. It will have to be justified by professional merit.”

PRO: ENSURING THE QUALITY OF PUBLIC SERVICES National salary, eligibility, and performance standards can ensure consistent quality of public services. By contrast, decentralized management of human resources might lead to deterioration of public services in localities that have little fiscal and managerial capacity to set competitive compensation packages that would attract qualified staff. Therefore a unitary hiring system might ensure that the necessary skills are present locally in all regions. However, this approach would also require centralized funding of the wage bill.

PRO: EQUITY OF PAY Decentralized determination of public sector pay raises the question of equity. This concerns the extent to which individual public sector employees feel “well paid,” both relative to other jobs (public or private) and in terms of their ability to afford a decent standard of living. Those OECD countries having implemented a uniform job classification system for the entire civil service (Denmark, Spain, and France) seem to secure equity and uniformity of pay. However, there is a trade-off between equity and merit. For example, in France the use of a single fixed multiplier implies equal treatment for all, while the general nature of the procedures for determining progress along the scale of index points reflects individual performance.

In Australia, one of the reasons for the return to servicewide bargaining during 1995–96 appears to have been the opposition of unions to the inequalities of pay. These inequalities were expected to arise under a system in which local pay raises were paid for by productivity gains where the opportunities to achieve productivity gains differed substantially among departments.

PRO: A UNIFIED LOCAL CIVIL SERVICE MAY CREATE NATIONAL COHESIVENESS For countries with many different ethnic groups and a geographically divided territory—such as Indonesia with over 4,000 islands and hundreds of ethnic groups, as well as many Sub-Saharan countries—it has been argued that a unified subnational civil service with central rules and controls for access, qualifications, and pay may act as a cohesive or unifying factor by creating access to jobs and positions for different people in different parts of the country.

* * *

Next we present some of the arguments that have been used against central government intervention and control of salary setting at the subnational level and in favor of discretion and autonomy in this matter by local governments.

CON: LIMITED MANAGERIAL FLEXIBILITY A countrywide pay scale does not offer management sufficient flexibility in meeting organizational requirements. Some room for variation in conditions of service and pay scales would allow subnational governments the opportunity to hire staff that matches local needs and budget constraints. The lack of flexibility detracts from the ability to keep more able employees or take advantage of the local labor market conditions. Flexibility allows regions to develop unique, cost-effective ways to attract qualified civil servants (for example, in-school housing for teachers). National wage scales limit the subnational government’s ability to compensate employees for varying costs of living or to attract people to particular jobs or locations.

One major advantage of pay decentralization is that it forces local management to develop a reward strategy that is consistent with an agency’s wider strategic goals. Indeed, delegating responsibility for establishing salaries to agencies may be viewed as a mechanism for stimulating agencies to develop an explicit and consistent set of strategic goals: to ask themselves what “business” they are in and how best they might serve their customers (Elliott and Bender 1997).

CON: NATIONAL PAY RAISES DESTABILIZE LOCAL BUDGETS In many countries, subnational governments' services tend to be more labor intensive (such as education, healthcare, and public works) than those of the central government. Hence, the wage bills of subnational governments are typically much larger in relation to their local economies than the central government's wage bill relative to GDP. Consequently the fiscal costs of pay increases are lower for the central government than for subnational governments, which means the central government is likely to underestimate these fiscal costs.

At the start of Russian reforms, wage increases for the total budgetary sphere ordered by the central government crippled local budgets because the refunding of these costs to local governments was delayed. The steep increase (90 percent) in wages mandated in the November 1991 Presidential decree for budgetary sphere workers was not funded by the Ministry of Finance (MOF) until March 1992. The subsequent round of wage increases decreed in January (45 percent) and August 1992 (50 percent) was funded in the 1992 budget subventions, but only for oblasts that, according to MOF, did not have a surplus.

In Lithuania, an increase in the minimum wage was the cause of a major dispute between the government and municipalities when, in 1998, the government increased the basic wage from Litas 95 to 105. Several municipalities—most conspicuously Vilnius—responded by refusing to remit the personal income tax (PIT) owed by its staff to the government until adequate compensation was received (World Bank 2002a).

India maintains a tradition of establishing pay commissions every 10 years to advise the central government on the determination of public service compensation. The recommendations of such commissions are applicable only to employees of the central government. Nevertheless, their recommendations have a major impact, via the demonstration effect, on subsequent pay raises by state governments and state public enterprises. As the central government's own wage bill does not constitute a large percentage of the overall general government expenditure commitments, it has the financial ability to raise wages even beyond the levels recommended by the pay commission. In fact, many states encounter significant fiscal difficulty in trying to keep up with central government pay (Rodden and Eskeland 2003).

CON: NATIONAL DECISIONS ON PAY ARE INSENSITIVE TO VARIATIONS IN LOCAL CONDITIONS Universal or countrywide pay scales detract from the necessary flexibility to take advantage of favorable labor market conditions or the necessity of offering higher compensation in the presence of labor shortages. In contrast, decentralized pay could be made up of many components. This would allow many diverse factors, such as pressures on the local or occupational job market, individual performance, and others to be taken into account.

In China, there is a tendency for nationally unified wage systems to set wages at levels closer to those of large, coastal urban areas (World Bank 2002b). These wages are very high relative to local levels in the poor regions. In Jishishan County (Gansu Province) they are said to be five to six times the local level. These high wages act as magnets for public employment and create great pressures on local governments to acquiesce to popular demands to absorb more employees. These national wages are sometimes far above local wage levels and local governments are unwilling or unable to pay them. Because salary levels are set nationally, local governments often are unable to meet nationally mandated increases. Authorities have reported that they cut salaries in such cases. In Qinghe County, for instance, officials have moved funds from the capital budget to cover wage increases.

CON: PERVERSE INCENTIVES A national pay scale combined with local hiring authority creates perverse incentives for local governments. In China, in times of uncertain and unstable transfers the only certainty local governments can count on is the commitment to pay salaries. It is no wonder, then, that budgetary expenditures are loaded up with wages and salaries, as

local governments move other (more discretionary) spending off their budgets. For poor regions, adding staff is often the only mechanism for getting more transfers from higher levels.

In India, based on the size of its state plan, each state is guaranteed a certain amount of financing, which is a mixture of grants, loans, and access to centrally guaranteed funding from financial institutions. The provision of central financing for the wage components of new programs for their first five years has also created a bias in favor of the expansion of programs and the ballooning of state civil services. Overstaffing in response to populist demands, along with the adoption of wage parity with the center by most states during the 1980s, has contributed to dwindling capital expenditure. Consequently, public investment at the state level has fallen far short of plan intentions and the resultant productivity of labor is low. During the late 1980s and early 1990s, as state resource constraints hardened, states maintained or increased their real expenditures on wages and salaries and subsidies, while reducing expenditures on socially productive capital and maintenance and permitting expenditures on basic healthcare and primary education to stagnate.

CON: UNDERMINED ACCOUNTABILITY TO THE CONSTITUENCY The lack of local governmental discretion over staffing and wages eventually makes local residents look to the central government for a solution of their problems (for instance, wage arrears). Like transfer dependence, central regulation of local fiscal affairs can send a strong signal to local politicians, voters, and creditors that the central government is ultimately responsible for local fiscal outcomes (Rodden and Eskeland 2003). This can be illustrated by the story of Brazil. Through much of the 1980s and 1990s, Brazilian states faced few effective hierarchical limitations on their spending or borrowing decisions. The only restrictions placed on the states actually had the opposite effect: by placing restrictions on states' ability to control their expenditures on public sector wages and pensions, the center bolstered the perception that states were unable to adjust.

International Experience

Countrywide wage scales for all public employees are common in unitary or centralized systems of government organization. However, this centralization of pay determination is mostly due to union demands. Most federal systems allow for considerable regional variance with respect to compensation levels, in consideration of standard-of-living distinctions, labor market differences, and recognition of subnational constitutional autonomy. Moreover, since the late 1980s and early 1990s, some unitary OECD countries have delegated authority for budget and staff management to local offices and introduced individual or collective performance-based pay. The idea that the performance of public sector workers can be enhanced by linking their pay to their performance has gained widespread acceptance in all OECD countries and especially in Australia, Sweden, and the United Kingdom. Moreover, performance pay has come to be viewed as a mechanism for both increasing public sector efficiency and improving the quality of public services (Elliot and Bender 1997).

The same level of centralization in the system of public pay determination can be structured very differently in different countries in order to take job market pressures for specific groups of professionals into account. In Germany, Spain, and France, collective bargaining on pay is centralized, and the same agreement applies to all civil servants regardless of the functional subsector or level of government in which they are employed. Different arrangements arise in connection with employment status, however. In Germany, civil servant pay is set by law. In Spain, senior civil servant pay is excluded from collective bargaining. In both countries, the pay for contractual workers (those without civil servant status) is determined at the department level.

In the Netherlands, a central pay agreement applied to all public sector employees until 1993. Since then, pension matters have remained centrally negotiated, but negotiations about the other terms of employment, including pay matters, are now organized separately for each

of the eight functional subsectors of public service (ministries, education, police, defense, justice, provinces, municipalities, and Polder Boards (water boards)). In Italy as well, pay negotiations are held in each of the eight functional subsectors defined as part the 1993 civil service reforms. The police and armed forces, university professors and other academic staff, judges and prosecutors, as well as senior civil servants are excluded from these negotiations.

Countries with a uniform job classification system for the entire civil service (Denmark, Spain, and France) can guarantee equity and uniformity of pay. However, these systems are rigid and cannot adapt to employment trends and job market pressures. Spain, with its specific supplements, and France, with its new index-related premium (*nouvelle bonification indiciaire*, or NBI), have attempted to respond to labor market realities. However, the costs of these measures, financed by the central government, contribute to the increase in the pay bill.

Other countries (such as Australia) have implemented a job classification system based on broad banding, in order to adapt their workforce to the needs of their various services. Yet, when agencies and departments are given the opportunity to implement their own classification systems (for example, the United Kingdom), salary increases in one agency may lead to increases in other agencies as well, undermining the effectiveness of decentralized pay determination.

Pay determination systems vary not only across countries but also within countries. In some cases the pay system of the civil service offers only a partial picture of pay determination in the extended public sector. In particular, this is the case for the United Kingdom, where pay reform focused on the civil service. At the same time, different systems based on collective bargaining or pay review bodies apply to the national health service, the education sector, law enforcement, and local authorities. In the United States and other federal countries, systems differ based on occupational groups, states, and localities.

4

Subnational Government Responsibilities

The assignment of expenditure responsibilities and the devolution of functional responsibilities to the regional or local government level is the first critical step in the design of a decentralized system of finance. Getting the assignment of responsibilities right at the local, regional, and central levels is a fundamental step in improving efficiency and fairness and in increasing the accountability of government to its citizens. Getting the assignment of expenditure responsibilities right should be seen as encompassing not only the devolution of expenditure responsibilities to local and regional authorities, but also as reaching a more balanced participation of the private sector, both in the production/delivery and in the financing of public services.

The Russian Federation has made significant progress since the early years of the transition in resolving some key expenditure assignment issues as part of its transition to a new federalism, but still has a long way to go in clearly defining the roles of the different levels of government in financing and providing public services. In particular, the role of governments at all levels and the public sector as a whole still needs to be optimized in the new market economy.

In theory it is rather simple to define the role of the public sector in a market economy: government intervention is only justified when the private sector cannot do a better job. That is, the operation of any activity should be left to the private sector if there is no economic argument (that is, ensuring efficiency or economic stability) or social rationale (that is, enhancing the equitable distribution of resources) for the government to intervene, regardless of other considerations. In this regard, apart from pursuing social equity objectives, the public sector intervention would be basically justified when there is a compelling market failure, such as those resulting from the presence of public goods (national defense, public order, and security), externalities (pollution, spillover effects), and uncontestable markets (public utilities, communications—where decreasing costs may drive out competitive forces). In addition, the government should pursue social equity objectives in a way that minimizes distortions on the working of the market economy. In practice, the delineation of the public sector in relation to the private sector is often much more complex because the process of devolution of functions and powers is eminently political in nature.

Difficulties in the implementation of expenditure assignments should always be expected, especially in the absence of a clearly established strategic decentralization policy and institutional mechanisms for cooperation and dialogue. From a practical side, the most formidable roadblock to the successful devolution of functional or expenditure responsibilities is the lack of administrative capacity at the local level. Successful implementation of fiscal decentralization policies often requires a strong capacity-building effort and always a certain level of political consensus, dialogue, and consultation. Decentralization policies that simply aim to push fiscal responsibilities down to subnational governments risk being ineffective and often end up adversely affecting the delivery of public services. The Russian Federation during the early years of the transition provides a textbook example of how not to tackle expenditure assignment reform. From 1992 to 1994, the federal government offloaded numerous federal expenditure responsibilities to the subnational governments, especially in the area of social welfare, with the clearest objective being to try to balance the federal budget. It took many years for legislation and practice to correct those hastily decided reassignments, which in practice became unfunded expenditure mandates to subnational governments.

The reform of expenditure assignments in the Russian Federation, perhaps reflecting the federal constitution of the country, always has focused a lot more on the division of responsibilities

between the federal and regional governments and much less on the division between the regional and local governments. The clarification of subnational government expenditure responsibilities is a more recent phenomenon. It is important to distinguish between the assignments as stated in the legislation and the *de facto* distribution of responsibilities.

This chapter reviews general principles and best practices on expenditure assignments and compares them to the actual arrangements in Russia. It includes the analysis of expenditure composition at the regional and local levels and addresses disparities in expenditure levels between different types of jurisdictions. In addition, it identifies main issues in expenditure assignment, reviews proposed changes, and suggests other options for policy reforms.

Principles and Best Practices Concerning Expenditure Assignments

There is no single fixed or established way in which expenditure responsibilities should be assigned among government levels. The extent and nature of decisionmaking power exercised by lower-level authorities varies considerably from country to country and it often changes from time to time in any particular country depending on changes in technology for the supply of public goods and services and preferences among citizens. In many unitary countries subnational governments are not constitutionally empowered to make independent decisions over a range of specified public functions or services. In federal countries, subnational governments tend to have more responsibilities assigned to them and more discretion to implement them.

However, there are better and worse expenditure assignments. This judgment on the advisability of assignments is based on some general principles from economic theory and international best practices. According to these general principles, policy decisions on economic stabilization and income distribution are best assigned to the central government, while some of those related to allocative efficiency (how to best use the resources available to provide goods and services) may be assigned to the local governments. Otherwise, when decisions on economic stabilization and income distribution are left to the local governments, wrong incentive and conflicts may arise, and policies may be ineffective and unsustainable. Observance of the principle of subsidiarity—that is, making decisions at the lowest possible level of government consistent with allocative efficiency, including the internalization of the benefits within the geographic area and the provision of services at minimum cost—is fundamental in defining the scope of local governments.

More specifically, the national government retains responsibility for the provision of national public goods, international affairs, monetary policy, regulation, redistributive policies including transfers to persons and businesses, fiscal policy coordination, regional equity, and preservation/promotion of the internal common market. The intermediate level (regional governments, if present) may be responsible for public services for which the benefits do not accrue to the entire national territory as a whole, but to more than one local government jurisdiction. For instance, regional-level expenditure responsibilities often include the provision of tertiary healthcare services, regional infrastructure, intermunicipal issues, and local government oversight. Local governments typically take the responsibility for a variety of local services, including primary and secondary education, local streets, the provision of utilities, and the provision of local services such as refuse collection. Of course, the actual assignment of expenditure responsibilities among different levels of government varies from country to country.

Beyond areas of exclusive functions and responsibilities that are assigned to each government level, the law should clearly specify competences on shared functions. When responsibilities are shared between two or more government levels, the central government should keep decisions on overall policies, setting standards, and auditing the use of central government funds, while intermediate governments may be given competence on overseeing implementation and local governments should be more directly involved in providing and delivering services (and infrastructure).

The division of responsibilities between regional and local governments should respond to the country's conditions and be dependent on considerations such as economies of scale (arising

from indivisibilities and specializations), economies of scope (arising from related activities), cost-benefit spillovers, proximity to beneficiaries, consumer preferences, and flexibility on expenditure priorities. In this regard, because of different local population size, distinct rural and urban characteristics, and diverse fiscal capacities, the assignment of responsibilities could be asymmetric, where larger local governments may end up having relatively larger responsibilities and a higher degree of autonomy.

Legal Assignment of Service Responsibilities to Local Governments

The division of service responsibilities among the different levels of government in Russia still suffers from great ambiguity, especially between regional and local levels. Yet, the lack of clarity is not due to a lack of legislation. Expenditure assignments have been addressed in a number of laws throughout the transition years. Instead, among others the lack of clarity is due to poor—often tautological—definitions used in the legislation. In addition, the current legislation lacks discussion on the division of responsibilities when different levels of government are responsible together for the delivery of a public service; the legislation fails to specifically assign responsibilities for regulation, financing, and the actual implementation (delivery or production) of the public service.

The first law on local self-government of the transition established the scope of competences for each tier and type of local government.¹ However, other than administrative autonomy and authority to regulate land use and merchandise, the law did not enumerate the specific functions of local governments. Article 71 of the new Constitution of 1993 establishes exclusive authority of the federal government over a number of areas such as defense, common market, railways, and telecommunications. However, in the Constitution most functions, including education, healthcare, and social protection, are assigned jointly to the federal and regional levels. As to the local level, the Constitution establishes the local government's autonomy in governing local affairs but does not enumerate any direct responsibilities for them with the exception of the protection of public order (Articles 130 through 133).

Subsequent legislation has not brought any clarity to the assignment of primary responsibility for the shared functions.² As pointed out above, most laws contain tautological clauses, stating, for example, that the federal government is responsible for funding federal programs and maintaining federal facilities. Articles 84–87 of the new *Budget Code* provide somewhat more specific assignments for local governments. However, the assignment in the *Budget Code* mostly covers financial responsibility, and it tends to be silent on who regulates and delivers these services.

The *Budget Code* establishes the exclusive responsibility of the federal government for national defense, space exploration, federal courts, financial support to regional governments, and official statistics. For the regional level, the *Budget Code* specifies exclusive responsibility for supporting mass media established by regional governments and for providing financial aid to local governments. Finally, for the local level the *Budget Code* enumerates municipal housing and utilities, construction and maintenance of local roads, waste utilization, transit, and earmarked subsidies to population.

Nevertheless, the *Budget Code* often employs for expenditure assignments the same tautological language found in the previous legislation. It states that local governments are responsible for maintaining and developing the institutions of education, healthcare, culture, and others that

1. *Federal Law (No. 1550-1) on Local Self-Government in the RSFSR* (July 6, 1991).

2. Article 6 of the 1995 *Law on the General Principles of Organization of Local Self-Government in the Russian Federation* attempts to determine the scope of local affairs. However, for the most part it employs a tautological language stating that local governments are responsible for maintaining and developing the institutions of education, healthcare, culture, and others that are *owned* by local governments. The only specific assignments for local governments include the protection of public order, fire protection, urban planning and land management, transit, construction and maintenance of local roads, heating fuel supply, waste utilization, funeral services, and maintenance of cemeteries.

are “owned by or put under the authority” of local governments. In this sense, the delineation of expenditure responsibilities continues to be centered on facilities rather than functions provided with those facilities. This is a trait inherited from the previous regime. Rather than being a caretaker of the physical facilities, a decentralized government should be empowered and encouraged to actively manage the delivery of local government services and seek ways to satisfy the needs of the constituency within the spheres of responsibilities assigned to the local level. At any rate, as a result of the focus on ownership, the division of responsibilities for the core expenditure functions between the regional and local levels of government continues to be determined to a large extent by the distribution of social assets between these two levels of government. And the distribution of the ownership of social assets differs considerably from region to region.

Finally, similar to other pieces of federal legislation, the *Budget Code* fails to accommodate expenditure assignment under the different local government structures established within the Russian regions. The 1991 law on local self-government at least attempted to define the scope of competences for each tier and type of local government. Unfortunately, these assignments were not exclusive and the list of competences was essentially repeated for each type of local government. By contrast, the *Budget Code* refers to a single local level.

The De Facto Distribution of Responsibilities among Levels of Government

To a large extent, the actual division of responsibilities among regional and local levels in Russia is not based on any law but rather is a product of historical development and practical compromises. During 1992–93, as we have already mentioned, the federal government offloaded major expenditure responsibilities from the federal to regional and local levels. These expenditures included capital investments in many areas, social welfare, and price subsidies for social goods such as food, medicine, local transportation, and public utilities (see table 4.1). The largest increase of subnational expenditures was on the national economy (transfers to enterprises) and social protection (transfers to individuals).

In partial defense of these federal actions, it is important to note that the “offloading” of expenditure responsibilities was closely linked to privatization. In the Soviet system, centrally planned enterprises were responsible for the provision of many basic public goods and services. Hence, before being privatized such enterprises maintained huge social assets: housing, daycare centers, hospitals, recreation facilities, and so on. In addition, enterprises were partially responsible for maintaining school buildings through a scheme of enterprise-school partnership. In the early 1990s, privatization was accompanied by a process of divestiture of enterprises’ social responsibilities. This process resulted in a significant transfer of social assets and the responsibility for their financing from productive enterprises to municipalities.

Enterprises faced a choice either to maintain the infrastructure and partially offset their tax liabilities from the housing maintenance tax (1.5 percent of enterprise turnover) by the amount of expenses incurred, or divest and pay the turnover tax in full.³ The process of divestiture has been very uneven and accompanied by threats and strategic positioning by enterprises and local governments. For example, in some cases a bankrupt enterprise disrupted the maintenance of its social assets, while at the same time the municipality where the enterprise was located refused to assume the responsibility for the abandoned social facilities. In some cases these types of behavior led to the closing of facilities that could no longer be afforded (lacking federal direct or indirect subsidies), such as daycare centers. In many others, infrastructure—such as housing and clinics—was let to deteriorate, leading to poorer quality of services and a much higher bill for the maintenance and replacement of the infrastructure.

In 1994, the federal government stopped the *ad hoc* shifting of expenditure responsibilities down to subnational governments. However, the federal center started introducing new programs

3. In 2001 the local turnover tax was replaced by a 5 percent surcharge on the federal tax on corporate income, which was again replaced by an entitlement to a fixed share of the federal CIT collections in 2002.

Table 4.1 Distribution of Expenditures by Function among Levels of Government, 1992–2002

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
<i>Administration and justice</i>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Federal budget	80.5	79.6	81.4	63.1	63.7	62.0	57.8	60.1	65.1	65.9	58.3
Consolidated subnational budget	19.5	20.4	18.6	36.9	36.3	38.0	42.2	39.9	34.9	34.1	41.7
<i>Regional budgets</i>	5.5	6.1	6.2	—	19.0	—	11.7	20.5	18.4	19.1	24.7
<i>Local budgets</i>	14.0	14.3	12.4	—	17.3	—	30.5	19.4	16.5	15.0	17.0
<i>Defense</i>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Federal budget	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Consolidated subnational budget	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
<i>Regional budgets</i>	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
<i>Local budgets</i>	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
<i>National economy</i>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Federal budget	80.8	48.6	30.4	26.3	27.5	24.0	4.7	12.6	14.6	19.2	12.8
Consolidated subnational budget	19.2	51.4	69.6	73.7	72.5	76.0	95.3	87.4	85.4	80.8	87.2
<i>Regional budgets</i>	8.7	24.6	34.8	35.4	32.4	—	17.7	36.6	39.0	43.8	47.5
<i>Local budgets</i>	10.5	26.8	34.9	38.2	40.1	—	77.6	50.8	46.3	37.0	39.7
<i>Education</i>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Federal budget	33.8	19.5	19.9	15.9	14.5	13.2	13.4	14.4	18.0	19.4	16.5
Consolidated subnational budget	66.2	80.5	80.1	84.1	85.5	86.8	86.6	85.6	82.0	80.6	83.5
<i>Regional budgets</i>	14.5	18.9	19.1	19.7	18.0	—	21.0	21.0	21.0	23.0	23.2
<i>Local budgets</i>	51.8	61.6	61.0	64.4	67.5	—	65.6	64.6	61.0	57.6	60.3

(Continues)

Table 4.1 Distribution of Expenditures by Function among Levels of Government, 1992–2002 (Continued)

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
<i>Culture and mass media</i>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Federal budget	51.2	34.8	39.6	31.6	34.7	16.5	16.5	21.2	28.3	28.3	24.9
Consolidated subnational budget	48.8	65.2	60.4	68.4	65.3	83.5	83.5	78.8	71.7	71.7	75.1
<i>Regional budgets</i>	21.2	30.2	28.9	30.8	29.3	—	38.9	38.8	36.2	38.4	40.1
<i>Local budgets</i>	27.6	35.0	31.5	37.6	36.0	—	44.6	40.0	35.5	33.3	35.0
<i>Health and phys. ed.</i>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Federal budget	11.3	10.5	11.8	9.9	10.2	11.8	8.8	9.9	11.2	12.4	9.6
Consolidated subnational budget	88.7	89.5	88.2	90.1	89.8	88.2	91.2	90.1	88.8	87.6	90.4
<i>Regional budgets</i>	28.5	30.2	33.4	33.4	33.1	—	39.0	38.9	40.0	41.7	45.1
<i>Local budgets</i>	60.2	59.3	54.8	56.7	56.8	—	52.2	51.2	48.8	45.9	45.3
<i>Social protection</i>	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Federal budget	71.8	52.0	50.3	18.5	31.2	41.3	56.6	53.5	53.1	52.0	74.5
Consolidated subnational budget	28.2	48.0	49.7	81.5	68.8	58.7	43.4	46.5	46.9	48.0	25.5
<i>Regional budgets</i>	20.4	34.9	34.9	31.1	26.8	—	23.4	26.7	29.5	30.7	15.4
<i>Local budgets</i>	7.8	13.2	14.8	50.3	42.0	—	20.1	19.8	17.3	17.3	10.0

—, Not available.

Source: Freinkman, Treisman, and Titov (1999) and Ministry of Finance of the Russian Federation.

while making subnational governments responsible for their funding. The issuance of two major federal mandates (payment of monthly child benefits in May 1995; and subsidies for payment on various goods and services by disabled persons in November 1995) resulted in a sharp increase in the local government's share in the expenditures on social protection. Moreover, local governments' involvement in social welfare was not confined to the provision of mandated payments to local residents. Provision of social services at the facilities that were divested from privatized enterprises involved significant subsidization from local budgets.

Because Soviet firms provided higher shares of labor compensation in the form of social benefits compared to other countries in Eastern and Central Europe, Russia inherited a more highly subsidized economy at the start of the transition. Some economists, for example Commander and Schankerman (1997), have argued that the divestiture of social assets by enterprises may have resulted in "under-consumption of merit goods arising from the lagged wage adjustment" after removing social benefits from the total labor compensation. This prompted a transitional system of income support in the form of subsidies to individuals (such as child allowances) and enterprises (for instance, for housing and utilities). However, the Russian experience illustrates that, if ill-designed, the income support program can have disastrous fiscal consequences. First, if income redistribution is not funded by the center, it is likely to cause greater inter-regional inequality as poor localities scare off businesses with the excessive fiscal burden of the program while fiscally rich localities (those with oil wealth, for example) exploit their fiscal advantage. Additional inefficiency results from neglecting the incentive issues in either the supply of or demand for the subsidized goods and services.

In Russia, the amount of subsidies to utility providers continues to be determined *after* enterprises have incurred operating losses attributed to price controls. Needless to say, under these arrangements, enterprises are discouraged from either cost containment or revenue mobilization. In order to phase out subsidies, local officials would have to introduce more effective utility regulations based on performance measures, which would require some administrative efforts. Carrying out the Housing and Utilities Reform, launched in 1990, is still a daunting task for Russian policymakers. The main goals of the reform are to eliminate cross-subsidization of one category of consumers by others, to expand cost recovery, and to challenge the monopoly power of network utility distributors.

During the Soviet era, rents were set below maintenance costs; investment capital was administratively allocated to local utility companies; and higher prices for industrial consumers cross-subsidized lower energy tariffs for households (Martinot 1997). Apart from bringing about price distortions, cross-subsidization distracts from the transparency of public-resource allocation and poorly serves social goals as the subsidized tariffs are enjoyed by all households regardless of their income level. Improved cost recovery is sought both through cost reduction and cost sharing with final consumers. Cost sharing implies targeted subsidies to low-income households instead of transfers to utility companies.

In practice, the regional government's decision to assign particular expenditure responsibilities to the local level is in part determined by the arrangements in place for the federal financing of subnational governments. Starting with year 2001, federal equalization grants, which are unconditional, have been complemented with transfers earmarked for fulfilling the two major federal mandates—payment of monthly child benefits and subsidies for payment on various goods and services by disabled persons. These earmarked grants partially replace the equalization grants as the latter used to account for the mandates' burden in the assessment of regions' need. In addition, the regional share of value-added tax (VAT) collections has been entirely allocated to the federal budget for subsequent redistribution across regions via the earmarked grants. As a result of this reshuffle, some regions have opted for the centralization of the responsibilities for these mandates to the regional level rather than passing the earmarked funds on to the local level.

In summary, local government service responsibilities are not clearly fixed in federal legislation and in practice they vary from region to region. Moreover, very often the scope of local government service responsibility in a particular region is not formally prescribed in the

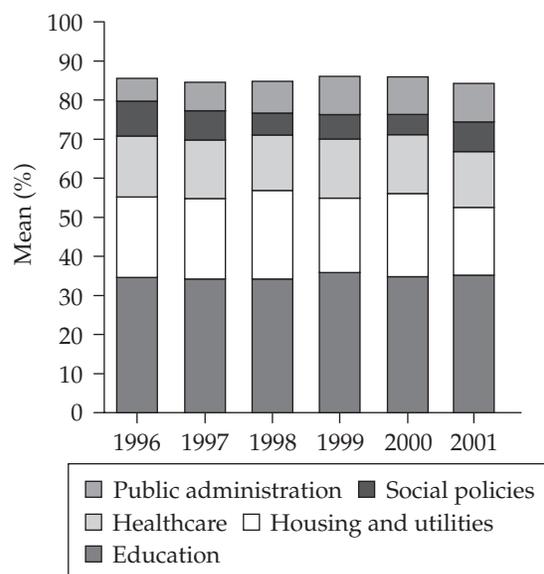
regional legislation but rather is a product of informal historical developments. There are several factors that determine the scope of local government functions in a particular region. First, local government responsibility depends foremost on the division of service responsibilities between the regional and local levels. Second, local government responsibilities depend on the scope of social services still provided by private enterprises to their workers. And third, local government responsibility depends on the relative role of extra-budgetary entities (for example, Health Insurance Funds) in the provision of social services.

Statistical Evidence on the Composition of Local Government Expenditures

Because local government service responsibilities are not explicitly outlined in either federal or local legislation, a natural way to assess the actual responsibilities is to analyze the actual expenditures of local governments along the lines of the functional budget classification. This is done in figure 4.1, which shows average values for the functional structure of local government total expenditures during 1996–2001.⁴ Across these years, about 70 percent of total local government expenditures is accounted for by three major functions: education (34.2–36 percent), housing and utilities (17.4–21.3 percent), and healthcare (14.1–15.7 percent).⁵ These shares have been fluctuating within the narrow ranges without any clear trend. The only time-trend revealed in the figure is a steady growth of the share for “public administration” in the total expenditures, from 5.7 percent in 1996 to 9.9 percent in 2001. It should be noted, however, that changes in employment of the local public administration tend to exhibit procyclical behavior, somewhat decreasing between 1996 and 1998 but starting to grow again with the economic expansion in 1999.

Although the shares of the three core functions in the total expenditures of local governments do not change over time, there are considerable variations between localities. Figure 4.2 juxtaposes

Figure 4.1 *The Distribution of Local Government Expenditure (All Types), 1996–2001*

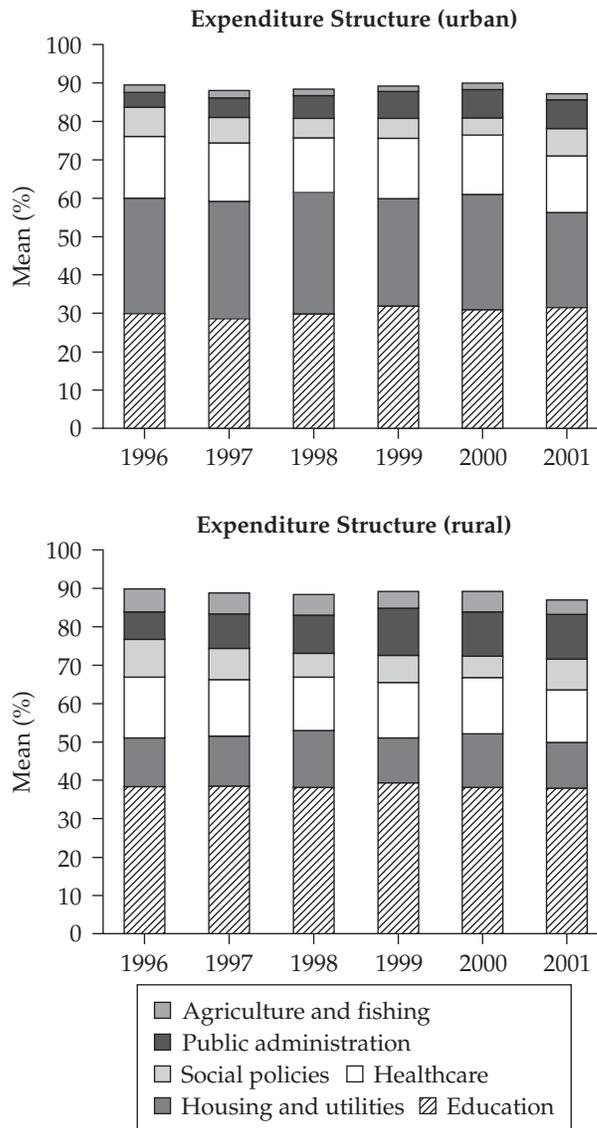


Source: Authors.

4. The average shares are computed on a representative sample of localities numbering from 693 to 1,945 in different years. In regions with two-tier local governments, data are aggregated at the top tier. The data set was collated by the Center for Fiscal Policy (Moscow).

5. Throughout this book, the total expenditures of local government exclude expenditures of earmarked funds even if the latter are eventually consolidated into the budget. The reason is to preserve comparability across years and also between those governments that incorporated the extra-budgetary funds in the budget and those that did not.

Figure 4.2 The Composition of Expenditures for Urban and Rural Local Governments



Source: Authors.

the average expenditure composition for urban and rural local governments.⁶ The share of expenditures on housing and utilities is twice as large for urban local governments as for rural jurisdictions. On the other hand, expenditures on education, public administration, and social protection account for a larger share of rural budgets than those for urban budgets. As expected, rural governments spend a larger share of their budgets on the support of agriculture but this is still at a modest level of 3–5 percent of their total expenditures.⁷

Additional evidence on the extent and ranges for variation in local government expenditure structure is provided in tables 4.2–4.4. There are cases involving localities whose expenditure structure is very different from the average profile. However, the shares of the three core

6. We classify a jurisdiction as “urban” if more than half of its population is urban.

7. Regional governments seem to play a larger role than do local authorities in supporting agricultural producers.

Table 4.2 *Distribution of Local Government Expenditure (All Types), 2001*

	<i>Mean</i>	<i>C.V.</i>	<i>Min.</i>	<i>Max.</i>
Public administration	9.9%	0.48	0.0%	76.2%
Law and order	0.9%	1.86	0.0%	51.5%
Industry, energy construction	4.0%	1.41	-12.3%	51.1%
Agriculture and fishing	2.8%	1.32	-8.6%	69.2%
Transport, roads, telecommunication	0.1%	6.66	0.0%	20.0%
Development of market infrastructure	0.7%	3.38	0.0%	66.8%
Housing and utilities	17.4%	0.71	0.0%	82.5%
Emergencies and natural calamities	0.2%	7.90	0.0%	54.7%
Education	35.1%	0.26	0.0%	65.5%
Culture, art, and cinema	4.2%	0.46	0.0%	14.0%
Mass media	0.3%	1.18	0.0%	3.5%
Health care and physical culture	14.1%	0.38	0.0%	38.2%
Social policies	7.7%	0.71	-0.1%	27.6%
Debt servicing	0.2%	8.18	-0.3%	83.4%
Other expenditures	2.3%	2.51	-48.7%	58.3%

C.V. Coefficient of variation.

Note: Negative expenditures arise from government lending, which is accounted for as positive expenditure at the time of lending and negative expenditure at the time of repayment.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

Table 4.3 *Distribution of Urban Local Government Expenditure, 2001*

	<i>Mean</i>	<i>C.V.</i>	<i>Min.</i>	<i>Max.</i>
Public administration	7.6%	0.56	0.0%	76.2%
Law and order	1.1%	1.15	0.0%	6.8%
Industry, energy construction	4.2%	1.36	0.0%	42.3%
Agriculture and fishing	1.5%	1.34	-0.1%	23.4%
Transport, roads, telecommunication	0.1%	4.93	0.0%	9.0%
Development of market infrastructure	1.0%	2.74	0.0%	66.8%
Housing and utilities	24.8%	0.46	0.0%	69.9%
Emergencies and natural calamities	0.2%	8.50	0.0%	54.7%
Education	31.5%	0.27	0.0%	65.5%
Culture, art, and cinema	3.2%	0.53	0.0%	11.6%
Mass media	0.3%	1.16	0.0%	3.5%
Healthcare and physical culture	14.7%	0.38	0.0%	38.2%
Social policies	7.1%	0.69	-0.1%	26.6%
Debt servicing	0.3%	10.07	-0.3%	83.4%
Other expenditures	2.3%	2.43	-11.0%	58.3%

C.V. Coefficient of variation.

Note: Negative expenditures arise from government lending, which is accounted for as positive expenditure at the time of lending and negative expenditure at the time of repayment.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

functions show the lowest coefficient of variation: 0.26 for education, 0.38 for healthcare, and 0.71 for housing and utilities. One way to interpret these figures is that there is consensus among local governments that education is more of a core function or an expenditure priority than are housing and utilities. However, as we noted above, a simpler interpretation is that the physical differences between rural and urban local governments are likely to cause more variation in housing expenditures than in educational spending.

Table 4.4 *Distribution of Rural Local Government Expenditure, 2001*

	<i>Mean</i>	<i>C.V.</i>	<i>Min.</i>	<i>Max.</i>
Public administration	11.6%	0.38	1.7%	39.9%
Law and order	0.8%	2.48	0.0%	51.5%
Industry, energy construction	3.8%	1.46	-12.3%	51.1%
Agriculture and fishing	3.7%	1.15	-8.6%	69.2%
Transport, roads, telecommunication	0.1%	7.31	0.0%	20.0%
Development of market infrastructure	0.5%	4.21	0.0%	43.8%
Housing and utilities	12.0%	0.85	0.0%	82.5%
Emergencies and natural calamities	0.2%	7.20	0.0%	39.4%
Education	37.8%	0.23	0.1%	64.6%
Culture, art, and cinema	4.9%	0.36	0.5%	14.0%
Mass media	0.3%	1.17	0.0%	2.8%
Healthcare and physical culture	13.7%	0.39	0.0%	33.8%
Social policies	8.0%	0.72	0.0%	27.6%
Debt servicing	0.2%	3.48	-0.3%	11.1%
Other expenditures	2.3%	2.58	-48.7%	44.9%

C.V. Coefficient of variation.

Note: Negative expenditures arise from government lending, which is accounted for as positive expenditure at the time of lending and negative expenditure at the time of repayment.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

Disparities in Expenditure Patterns at the Local Level

The evidence presented above suggests that the composition of local government expenditures might vary depending on the "need" for a particular service in different localities. Accordingly, urban governments have inherited the responsibility for centralized heating while rural governments have to provide veterinary services. However, besides the scope of assigned responsibilities and the local need for these services, actual public expenditures in a locality depend on two additional components: the quality of public services selected by the local authorities, and the costs of public services in that particular locality (per unit of service quality). The combined impact of all three determinants of actual expenditures (need, quality, and cost) leads to a great disparity in actual per capita outlays on the same service by different localities.

To illustrate this disparity, we present the results of our analysis of local government per capita expenditures on healthcare. Appendix table 4.1.3 reveals that even in a homogeneous sample of 29 large cities, per capita outlays on healthcare vary by a factor of 2.7 (as indicated by the last column). To better understand the sources of variation in expenditures per capita, we attempted to isolate the component of this variation that is due to the factors outside local government control (including the extent of assigned responsibilities and the unit cost of service provision at an average efficiency level).⁸ Thus, for our sample of urban cities we estimated the objective "expenditure need," defined as the amount of money local governments must spend to achieve a given quality of public services given the exogenous factors they face.

The methodology and steps employed to objectively quantify local expenditure needs are reported in appendix 4.1. Essentially, the expenditure need for a particular city is derived from the average per capita regional-local expenditure on healthcare (RUR 412,600 in 1997) by applying adjustment coefficients to account for objective differences in service responsibilities, service need, and production costs. This resulting per capita amount is reported in the penultimate column of appendix table 4.1.3. The variation in the assessed expenditure need is smaller than

8. If the local government operates inefficiently, for example using the wrong combination of inputs of production such as labor and capital, the costs per unit of service will be higher than they need to be.

the variation in actual expenditure. Thus, the coefficient of variation in expenditure need is 0.16, down from 0.31 for actual expenditure. At the same time the ratio of maximum to minimum expenditure need is 1.97, compared to 2.71 for the actual expenditures.

These results show that objective expenditure need explains only a fraction of the disparity in healthcare expenditures. The detailed analysis suggests that affluence also plays a role so that wealthier cities provide higher levels of healthcare for the same need as determined by demographics. The residual variation can be attributed to differences in cost efficiency.

Main Issues in Subnational Expenditure Assignments

The assignment of expenditure responsibilities is likely to be the most difficult element of a decentralized system to get right. The problem is not only that the set-up can be easily politicized but also that it is technically difficult to properly address all the different aspects that go into a good assignment of expenditure responsibilities. The determination of what constitutes a good expenditure assignment is also likely to change over time. Therefore it is useful to remember here again that there is no such thing as a unique and perfect assignment of expenditure responsibilities that fits all. Expenditure assignments need to be flexible and adapt to the historical and idiosyncratic features of a country. Nevertheless, there are many assignments of responsibilities that can be considered inferior because they fail to comply with basic principles. The assignments of expenditure responsibilities at the subnational level in the Russian Federation fall into this latter category. The main issues that Russia needs to address in this regard include the following:

- *Overly broad and vague assignment of responsibilities and competences in the legislation.* Federal legislation has fallen short so far in assigning specific detailed authorities exclusively to specific government levels. Similarly, legislation falls short in clearly defining respective competences in the case of shared responsibilities. The system is still exposed to confusion and arbitrariness, since the competences (on setting policies, regulation, financing, and service delivery) have never been clearly understood, with responsibilities being disputed by different authorities and in some cases in practice being shifted back and forth.
- *Delineation of responsibilities still centered around facilities rather than public services provided by these facilities.* Consequently local governments are made responsible for maintaining and developing the institutions *owned by or put under the authority of* local governments. This is a remnant feature of planned socialism where the center makes a strategic decision regarding the services to be provided while local agents merely implement the centralized plan. As noted earlier, rather than being a passive caretaker of the entrusted facilities, a subnational government in a decentralized system should actively seek ways to satisfy the needs of its constituencies within the spheres of responsibilities assigned to it.
- *Decentralization of inappropriate functions.* In Russia both the regional and local governments have been assigned responsibilities that affect income redistribution and in some cases economic stabilization. In particular, local governments appear to carry out the bulk of social welfare programs. Theoretically at least, it might be the case that local governments have informational and comparative advantages in targeting the needy, but local governments should not have the responsibility for financing the bulk of these programs. But even the argument on the informational advantage of local governments does not seem applicable to the Russian case, since regional governments have a network of social security offices in every locality (so called *SOBESes*). The fact that many regions have recentralized the provision of child benefits after the federal government started funding it with earmarked subventions only proves the case that

certain governments are decentralized to the local government level for political convenience rather than for efficiency reasons.

- *Fusion between local and federal-regional sectors.* Although the new Constitution insulated local self-governance from the sovereignty of the state, there is still a significant amount of fusion between state (federal-regional) and local sectors. For example, although the social rights, such as access to education and healthcare, are guaranteed in the federal Constitution, most of the social infrastructure was transferred to local governments in the course of divestiture from privatized enterprises. As a result, the implementation of many federal and regional policies continues to rely on local authorities. However, the new legislation limits state control over activities of local self-government, with the latter defined as the management of local affairs by local residents under their own responsibility. Bringing about effective governance will require coordination among the levels of government based on mutual agreement and clearly outlined accountability for delegated tasks.
- *Unequal administrative and technical capacities.* The unequal administrative and technical capacities of local governments are a factor that has contributed significantly to the weak implementation of policies and has prevented the efficient delivery of quality services in numerous local government jurisdictions. While these capacities may exist at an adequate level in larger urban centers, this does not seem to be the case in rural areas. There is a tremendous difference in terms of local government capacity between urban and rural administrative units, especially at the lowest level of the government. Thus, concentration of population in urban areas makes it economically desirable for the local sector to provide a wider range of services that involve significant indivisibilities (museums, zoos, theaters, and the like). Also, residents of urban areas may demand different types of public services than do rural residents. For example, urban governments allocate a significant share of their resources to intracity transportation and public housing and utilities. More to the point, federal tax legislation skews the distribution of the local tax base toward urban areas. This results in considerable disparities in per capita local expenditures between rural and urban jurisdictions.
- *Perverse incentives.* The current system of subnational finance in the Russian Federation contains perverse incentives that induce local government to use scarce public resources inefficiently. For instance, regional governments still interfere excessively in areas for which exclusive responsibility already has been assigned to the local government level. Regional governments do so by controlling and approving tariff policies of utilities, determining local wage rates, and imposing local employment “norms.” These interventions distort local resource allocations and create inefficiencies in the production and delivery of public services (including social services and public utilities).

The Kozak Commission’s Proposed Changes in the Delineation of Expenditure Functions

The Kozak Commission recognizes the lack of a clear assignment of expenditure responsibilities at the regional and local levels of government and—among other recommendations—attempts to establish a clear delineation among all levels of government (table 4.5). Out of the functions assigned by the Constitution jointly to the federal and regional levels, the Kozak proposals enumerate those that have to be independently carried out by regional governments at their own expense. Although these functions can be subject to general regulation by the federal government, the level and procedures for delivering them cannot be mandated. For functions remaining as joint federal-regional responsibilities, the Commission suggests that the federal government would have to provide earmarked funds and detailed regulations for regional governments.

Table 4.5 Expenditure Responsibilities of Regional and Local Government According to the Kozak Proposals

<i>Expenditure functions</i>	<i>Settlements</i>	<i>Urban districts</i>	<i>Municipal districts</i>	<i>Regions</i>
<i>General public services</i>				
Archives	Formation of the settlement archive	Formation and depository of the municipal archive	Formation of the district archive, depository for district and settlement archives	Formation and depository of the regional archive
<i>Public order and safety</i>				
Street patrolling	None	By local police	By local police	None
<i>Education</i>				
Preschool	None	Maintenance of buildings, financing of utilities, and organization of the rest with regional financing	Maintenance of buildings, financing of utilities, and organization of the rest with regional financing	Financing of wages, textbooks, equipment and supplies
Primary	None			
Secondary	None			
Tertiary	None			
Vocational/technical	None	None	None	Provision
<i>Health</i>				
Ambulance	None	Organization of ambulance service, primary healthcare, and obstetrical service	Organization of ambulance service, primary healthcare, and obstetrical service	Airlift
Outpatient clinics and hospitals	None			Provision of special treatment (dermatology, TSD, tuberculosis, narcology, oncology, and so on)
Public health	None	None	None	
Insurance for the unemployed	None	None	None	Financing
<i>Social security and welfare</i>				
Social protection	None	None	None	Child benefits, support to labor veterans, orphans, homeless/destitute children, single mothers, and mothers with many children
Guardianship	Facilitation	Provision	Provision	None
Labor market policy	None	None	None	Retraining, public works

Table 4.5 Expenditure Responsibilities of Regional and Local Government According to the Kozak Proposals (Continued)

<i>Expenditure functions</i>	<i>Settlements</i>	<i>Urban districts</i>	<i>Municipal districts</i>	<i>Regions</i>
<i>Housing and community amenities</i>				
Social housing	Provision	None	None	Financing of subsidies to households
Garbage/waste	Collection	Collection and utilization	Utilization	None
Water	Organization	Organization	None	None
Sewerage	Organization	Organization	None	None
Planting of greenery, forest protection	Provision	Provision	None	None
Street sweeping, lighting, and signs	Provision	Provision	None	None
Funeral services and maintenance of cemeteries	Organization	Organization	Organization	None
Firefighting/fire protection	Basic fire protection	Basic fire protection	None	Organization of firefighting
<i>Recreational and cultural affairs</i>				
Physical culture development	Facilitation	Facilitation	None	Organization of regional programs
Libraries	Municipal libraries	Municipal libraries	Library collectors	Regional libraries
Museums	None	None	None	Organization and support of regional museums, folk crafts, other cultural activities
Arts	None	None	None	
Entertainment	Facilitation of service provision	Facilitation of service provision	None	None
Places of recreation	Arrangement	Arrangement	None	None
Protection of historical sites	Within the settlement	Within the settlement	None	Sites of regional importance
<i>Fuel and energy</i>				
Central heating	Organization	Organization	None	None
Heating fuel (natural gas, firewood, coal)	Organization	Organization	Organization (natural gas)	None
Electric power supply and lighting	Organization	Organization	Organization	None

(Continues)

Table 4.5 Expenditure Responsibilities of Regional and Local Government According to the Kozak Proposals (Continued)

<i>Expenditure functions</i>	<i>Settlements</i>	<i>Urban districts</i>	<i>Municipal districts</i>	<i>Regions</i>
<i>Agriculture, forestry, fishing, and hunting</i>	None	None	None	<i>Support to agricultural producers</i>
<i>Transportation and communication</i>				
Telecommunications	Facilitation of service provision	Facilitation of service provision	Facilitation of service provision	None
Road network	Organization within the settlement	Organization within the settlement	Organization between settlements	Organization between districts
Public transport	Organization within the settlement	Organization within the settlement	Organization between settlements	Organization between districts
Railroads	None	None	None	Organization between districts and suburban
Airlines	None	None	None	Organization between districts and suburban
<i>Other economic affairs and services</i>				
Urban planning and land management	Development of plans and zoning of settlement land	Development plans and zoning, maintenance of land records	Maintenance of land records, zoning of intersettlement lands	Planning the use of agricultural land
<i>Other expenditures</i>				
Environmental protection	Organization within the settlement	Organization within the settlement	Organization between settlements	Organization between districts
Prevention and handling of emergency situations	Participation in responding to local incidents	Participation on local incidents	Participation on district incidents	Regional and intersettlement incidents
Ecological control	None	Provision	Provision	None
Equalization of constituent jurisdictions	None	None	Equalization of settlements with district funds according to regional methodology	Equalization of districts and settlements according to federal guidelines

Source: Federal laws on *General Principles of Organization of Legislative (Representative) and Executive State Power Bodies of Subjects of the Federation* (No. 184-FZ of October 6, 1999; last amended July 4, 2003) and on *General Principles of Organization of Local Self-Government in the Russian Federation* (No. 131-FZ of October 6, 2003).

Regional governments would be solely responsible for providing support to some disadvantaged groups of the population, as well as maintaining an active labor market policy and providing support to agriculture. A majority of the remaining regional government responsibilities are to be shared with the local governments.

Several attempts are made in the Kozak Commission's proposals to specify how regional and local governments would complement each other in carrying out these concurrent functions. Some functions would be carried out on a territorial basis (such as the building and maintaining of roads, organization of public transportation, prevention and liquidation of emergencies, and environmental protection). For example, local governments would be responsible for building and maintaining roads within municipalities while the regional government would be responsible for building and maintaining roads between districts. For other responsibilities, including education, healthcare, and land regulation, the joint responsibility is identified by the type of service (for example, primary versus tertiary healthcare). However, within general education, responsibility is shared by expenditure item (buildings and utilities versus the rest). Responsibility for other services is shared on the basis of ownership of the facility: libraries, historical sites, sport facilities, public archives and so on.

The Kozak Commission's proposals also provide separate assignments at the local level. The proposed law assigns expenditure responsibilities separately to each of the (proposed) tiers of local government. These local government responsibilities could be delegated from one tier of local government to the other based on a contract basis, by specifying the terms, modalities for termination, calculation of compensation, and financial sanctions.

Settlement governments would be solely responsible for providing housing to the low-income population, basic fire protection, street lighting and signs, planting greenery and protecting forests, protecting historical and cultural memorials, facilitating entertainment services, facilitating physical culture development, and maintaining recreational sites.

Urban districts would be solely responsible for organizing the protection of public order by the municipal police; organizing general education; organizing ambulance, primary healthcare, and obstetrical services; ecological control and environmental protection; disposal of household and industrial waste; maintenance of land records; and safekeeping of the municipal archives. The municipal districts have the same responsibilities as urban districts with an additional responsibility for the equalization of constituent localities using district budget resources.

Responsibilities for other services would be assigned jointly to the lower and higher tier of local government. These concurrent responsibilities include organizing utilities provision, building and maintaining roads, organizing public transportation, organizing library services, administering guardianship over minors and the incapacitated, safekeeping of the municipal archives, regulating land use participating in the prevention and handling of local emergencies, and organizing funeral services and maintaining cemeteries. For these concurrent services, the delineation of responsibility is prescribed based on the territorial principle.

As proposed by the Kozak Commission, neither federal nor regional laws would be able to prescribe expenditure levels for the exclusive functions of local governments. But unlike in the case of regional government functions, this clause does not prohibit higher-level government to mandate forms for service delivery at the local level. The proposal also introduces a legal framework for federal and regional mandates imposed on local governments. In particular, article 63 of the draft laws based on the Kozak proposal requires that there be a special compensatory fund created in the regional budget that provides resources for financing federal and regional functions delegated to local governments. Federal and regional budget laws should determine every year the sizes of these funds for the planning year. The revenue sources of the regional fund are subventions from the federal compensatory fund and other regional revenue sources.

Overall, the Kozak Commission's proposals clearly represent a significant improvement over the ambiguous delineation of responsibilities outlined in the Constitution and more recently in the *Budget Code*. It is in particular quite important from the viewpoint of subnational

government finance that the proposed legislation enumerate exclusive responsibilities of sub-national governments other than just managing their own operations and property.

Compared to the present *de jure* and *de facto* division of responsibilities among levels of government in the Russian Federation, the Kozak proposal introduces a number of drastic positive changes. First, the financing of health insurance for the unemployed and most outlays on general education have been reassigned from the local to the regional level. Second, it appears that the current federal mandates on support to some disadvantaged categories of population would be reassigned as exclusive responsibilities of the regional governments. While the *Budget Code* currently assigns the responsibility for supporting the railway and air transportation sectors to the federal government, the Kozak proposal makes regional governments responsible for the organization of intermunicipal transportation by rail and air. Support for agricultural producers would be made a *de jure* exclusive responsibility of regional governments. Until now, this has been already a *de facto* responsibility of regional governments. While the *Budget Code* currently assigns firefighting and law enforcement services to be jointly financed by the federal and regional governments, the Kozak proposal makes firefighting an exclusive responsibility of regional governments. The proposal also separates municipal street patrols from federal detective services. Thus, under the proposal, the regional level essentially would be stripped of any law enforcement powers.

For the concurrent expenditure functions, the Kozak proposal falls short of explicitly specifying the “attributes” for each shared responsibility among two or more levels of government. The proposal is silent about who has the obligation for regulating, financing, and implementing these shared services. Instead, the proposal still employs terms, with no definition provided, such as “organization of service delivery,” “facilitation of service provision,” “participation in service delivery,” and so on. One suspects that this fuzzy language stems from the still-undefined scope of government. Thus, for example, “organization of utilities provision” may be read as “regulation” in localities that have full-cost recovery in the provision of utilities, but it may mean “funding” or even “delivery” in localities where this service is heavily subsidized or actually provided by the local government. While the reasons may be understandable, these situations could be better handled in transitory clauses in an appendix as opposed to using fuzzy language in the main text of the expenditure assignments.

Besides increasing clarity in the assignment of responsibilities, the Kozak proposal improves the delineation of expenditure responsibilities in other respects. First, it makes clear the principle that for “delegated” (as opposed to “own”) responsibilities, the upper-level government delegating those responsibilities has the power to regulate but also the obligation to ensure there is sufficient funding to provide those services. In particular, expenditure mandates from above require the provision of explicit budget funds every year for the governments issuing the mandate. Second, it makes clear that for all other local responsibilities (own or nonmandated), upper-level governments cannot introduce expenditure norms. These provisions should go a long way toward smoothing intergovernmental relations among all levels of government in the Russian Federation.

By making regional governments responsible for financing most outlays on general education, the proposals resolve the fusion between local and federal-regional sectors. However, the complete centralization to the regional level is problematic because school buildings were transferred to localities in the early 1990s. On the other hand, partial responsibility of local governments over general education presents a danger of suboptimal decisions especially regarding the much-needed rationalization of the school network. For example, there is a need for establishing clear rules for making joint regional-local decisions over the opening and closing of schools in localities.

Options for Further Policy Reforms

As we just noted, the Kozak proposal for intergovernmental reform and expenditure assignments addresses a number of difficult issues in this area that had been previously ignored. If

approved into law, the proposal will significantly improve intergovernmental fiscal relations in Russia. However, there may still be a need for further reforms and improvements on what has been recommended by the Kozak Commission. Here are some of the issues that may require further consideration:

- *Decentralizing adequate decisionmaking powers on public expenditures asymmetrically when necessary.* As discussed above, there are significant differences in terms of local government capacity between urban and rural administrative units, especially at the lowest level of government. By introducing two levels of government at the local level and assigning different expenditure responsibilities to them, the Kozak proposal recognizes the different capacities of local governments and in a way provides for their asymmetric treatment. The question is whether two packages of expenditure assignments at the local level are adequate to recognize the differences in administrative capacity among local governments. An argument could be made for allowing further differentiation of expenditure responsibilities to adapt them to the different administrative capacities at each level of local government. Thus, regional governments may be enabled to assign a reduced package of responsibilities to those local governments that lack technical and administrative capacities.
- *Clarifying shared expenditure responsibilities among different levels of government.* We have seen that for concurrent or shared functions, the Kozak proposal falls short of explicitly specifying the “attributes” for each shared responsibility among two or more levels of government. Functional responsibilities can be further clarified by specifying in the law what level of government is responsible for regulating, financing, and implementing each of the shared responsibilities and to provide clear (non-tautological) legal definitions for these concepts. This should also help eliminate the fuzzy terms still used in the Kozak proposal. However, there is no guarantee that the assignment of specific attributes (on regulation, financing, and delivery) for the existing programs would completely clarify responsibilities. For example, it is possible for regulation powers to be attributed to two levels of government. In such cases, there should be further clarification in the law on what particular power of regulation the different levels of government will have.
- *Developing institutions for cooperation and dialogue.* There is no formally legislated expenditure assignment that can be exhaustively complete. There will always be important details as well as new situations that will need to get resolved in practice. Solving the differences in perspective from the two levels of government should not be a one-way street, with the upper-level governments always dictating solutions. The effectiveness of expenditure assignments and the day-to-day operations of the intergovernmental system could be much improved by the creation or further enhancement of institutions and mechanisms for coordination and dialogue among different levels of government in all sorts of expenditure assignment issues. Those responsible for general education at the regional and local levels should therefore meet periodically and establish open channels of communication in order to resolve all matters of interpretation and uncertainty in their concurrent responsibilities. There are also significant possibilities for horizontal cooperation among local government units for the joint provision of services. Interjurisdictional cooperation could occur through public or private companies (or an association among them) in order to explore economies of scale (by merging or associating existing companies), or to internalize existing externalities among subnational jurisdictions.
- *Defining a new strategic direction for the provision of public services.* Reforms of expenditure assignments could aim at making the cost of public services (including utilities) more affordable and sustainable in the long term by, among other things, continuing to strengthen the role of the private sector in the financing and delivery of public services.

Dual tariff mechanisms still practiced by some local governments and utility companies should be eventually eliminated to reduce inefficiencies and distortions. Subsidies for utilities, housing, transportation, and others should be directly targeted to the needy while tariff policies should be based on the principle of full cost recovery. Most local governments in Russia are running behind in the privatization of public service delivery. More privatization should help build a more efficient public sector and should enable it to provide better quality and cheaper services to citizens. Increasing private sector participation could be an important goal to keep on the front agenda of subnational government sectoral policy reform, such as in the case of housing, where there is still widespread reliance on public sector provision.

- *Promoting administrative and technical capacity building.* Some important issues in expenditure assignments arise because many subnational governments do not have adequate administrative capacity and because this capacity differs markedly among local governments. Both federal and regional governments should promote administrative and technical capacity building, including interpretation of the legislation and general regulations.

Appendix 4.1. A Methodology for Measuring Cost Disparities across Local Governments

The goal of this appendix is to demonstrate the development of a methodology that would allow us to estimate the impact of input prices and a variety of environmental cost factors on healthcare expenditures, while controlling for the division (actual assignment) of responsibilities and the variables that determine service level.

We explicitly distinguish between the quality of public services chosen by local authorities and their expenditures toward the provision of these services. The quality of public services will be proxied by service level. This is because we believe that, although far from ideal, the endogenous quality of public services can be captured with measurable quantities of public output—for example, the number of classes given or the number of patients consulted. Obviously, by choosing the volume of the public outputs produced, local governments can affect the resulting quality of these public services. By contrast, we can estimate the relationship between the client-group size (for instance, school-age children) and the public output (for instance, the number of patients treated.) The estimated relationship will allow us to single out the variation in public outputs that is due to the exogenous environment and discard the variation that is due to local government choice on the quality of public services.

In the second stage, we estimate the average cost function for producing public outputs. The estimated coefficients for the explanatory variables can serve as weights accounting for the impact of each of these cost factors on local government expenditure needs. Thus, we assume that there is some well-defined production structure for the public outputs, characterized by a well-behaved cost function translating output levels and input prices into some objective amount of costs. At this point, it is irrelevant whether we estimate the “average” or “best practice” technology. It has been shown in the econometric literature that under either assumption, the slope parameters can be consistently estimated by ordinary least squares.⁹ Thus, we can safely fit the cost function through the middle of a cloud of points plotted for each of the actual cost outcomes in our sample. However, if we are interested in the nature of deviations from this “average” technology and want to distinguish between symmetric random shocks and systematic slack or inefficiency in the production of services, we would have to apply a special treatment to the estimated residuals.

9. For example, see Greene (1993).

In formal terms, we consider the following relationship:

$$E_i = E(y_{1i} \dots y_{Ni}, w_{1i} \dots w_{Mi}, c_{1i} \dots c_{Li}) + e_i$$

Here, y_i 's are outputs, w_i 's are input prices, c_i 's are environmental cost factors, E_i is the actual expenditure toward producing the y_i 's, and e_i is the deviation from the "true" cost function in locality i . Furthermore, we do not rule out that $e_i = v_i + u_i$, where $u_i \geq 0$ stands for additional costs due to slack, while v_i is unrestricted and captures unobservable cost shocks that are outside local government control.

For the sake of simplicity,¹⁰ we specify the cost function in the log-linear form:

$$\ln E_i = \alpha + \sum_{n=1}^N \beta_n \ln y_{ni} + \sum_{m=1}^M \delta_m \ln w_{mi} + \sum_{l=1}^L \gamma_l \ln c_{li}$$

First, we estimate the relationship between the client-group size, as determined by the demographic structure, and the public outputs: bed days in hospitals, visits to the hospital doctor,¹¹ number of patients treated in hospitals, and visits to the polyclinic doctor. The estimated relationship allows us to single out the variation in public outputs that is due to the exogenous environment and discard the variation that is due to local government choice on the quality of public services. We explicitly account for the per capita amount of available fiscal resources to control for income or pecuniary effects.

The marginal impacts of differences in the size of various client groups on the observed level of healthcare production are reported in appendix table 4.1.1. Overall, the regressions perform relatively well, managing to explain almost 70 percent of variation in the hospital outputs and about 40 percent of variation in the polyclinic outputs. As expected, the level of healthcare outputs provided by hospitals is positively related to the availability of fiscal resources. After controlling for this pecuniary effect, we estimate that children between seven and nine years of age require more bed-days in hospitals than the benchmark group (working-age population). On the contrary, children aged 0–6 and 10–15 seem to require less bed-days in hospitals than the benchmark group. The impact of the high school-age population and pensioners is not statistically different from that of the benchmark group. For the polyclinics, a larger share of the middle school-age population and pensioners is associated with fewer visits to the doctor.¹²

Both in the hospitals and polyclinics, all types of healthcare outputs are negatively related to the employment of the rural population. As expected, the particular organization of outpatient care in rural areas leads to a positive relationship between the share of the rural population and the number of visits to the hospital doctor.

10. For a wide range of production functions, the cost function can be expressed in the translog form—Cobb-Douglas, translog, generalized Leontief, generalized Cobb-Douglas, and some others. We restrict the translog to a log-linear form so that we would have a single weight for every cost factor. Thus we implicitly assume constant elasticity with respect to each cost factor. However, the interaction and quadratic terms can be included if we estimate the cost function for other purposes, such as measuring inefficiency. In this case, the elasticities would be a function of the explanatory variables and thus could be evaluated on particular values (for instance, sample averages).

11. We observe visits to the hospital doctor rather than to the polyclinic doctor mostly in rural areas. This is because in rural areas outpatient services are provided at the Central Rayon Hospital (CRH) rather than at a polyclinic.

12. Unlike the working-age population, pensioners are not motivated to visit a polyclinic just to get a sick-leave authorization.

The elasticities reported in appendix table 4.1.1 allow us to assess the relative need to produce healthcare service stemming from differences in the client group size. To determine the ultimate expenditure need, we also have to estimate relative cost differences in producing a unit of healthcare services. Appendix table 4.1.2 reports elasticity of the healthcare production costs with respect to output levels and input prices. The elasticity of costs with respect to each type of output is positive, ranging from 0.077 for visits to the polyclinic doctor to 0.256 for the number of patients treated in hospitals.

Average wage is the only information we have to account for differences in input prices.¹³ Out of the choice of wage measures averaged over the manufacturing sector, agriculture, and the total local economy, only the first one turns out to have a statistically significant impact on healthcare costs. The estimated cost elasticity with respect to wages in manufacturing is 0.134. Being the only employed indicator for input prices, the average wage is likely to account for the impact of differences in the general price level in the locality, and the estimated elasticity should have a broad interpretation.

We also attempted to account for the portion of costs covered through the Compulsory Health Insurance Program. Our explanatory variables include the local government coverage (in percentage points) of the total wage-bill of hospitals and polyclinics. As expected, this indicator is positively related to local government costs of healthcare provision. None of the environmental factors—such as urbanization or employment structure—appears to have a statistically significant impact on healthcare costs.

Appendix Table 4.1.1 *The Impact of Client Group Size on Output Level, 1997*

	<i>Hospitals</i>		<i>Polyclinics</i>	
	<i>Bed-days</i>	<i>Visits to the doctor</i>	<i>Number of treated patients</i>	<i>Visits to the doctor</i>
Share of pop. age 0–6	-1.024	-0.451	-0.184	-0.266
Share of pop. age 7–9	1.923	0.581	0.675	-0.510
Share of pop. age 10–15	-0.210	0.110	0.024	-0.676
Share of pop. age 16–17	-0.323	-0.126	-0.190	-0.166
Share of registered pensioners	0.054	-0.018	0.105	-0.490
Share of rural population	—	0.160	n.i.	n.i.
Share of rural population employed	-0.118	-0.096	-0.152	-0.462
Fiscal resources	0.296	n.i.	0.400	n.i.
R ²	0.68	0.66	0.66	0.42
Number of observations	232	232	232	302
Number of regions	12	12	12	13

n.i. Not included.

R². Coefficient of determination.

Note: Variables are expressed in logarithm—the marginal impact is the percentage impact of a 1 percent difference in the variable; region dummies are employed; the bold font indicates statistical significance at 10 percent and higher.

Source: Authors.

13. The price of the basic foodstuffs bundle (the subsistence level) is reported only for major cities.

Appendix Table 4.1.2 Estimation of the Healthcare Cost Function, 1997

	Variant 1	Variant 2	Variant 3
<i>Outputs</i>			
Bed-days in hospitals	0.188	0.309	0.122
Visits to the doctor in hospitals	0.244	0.440	0.268
Number of treated patients in hospitals	0.256	0.077	0.202
Visits to the doctor in polyclinics	0.077	0.062	0.098
<i>Inputs</i>			
Government-funded share of the wage-bill	0.249	0.272	0.443
Average wage	n.i.	0.061	n.i.
Average wage in manufacturing	0.134	n.i.	n.i.
Average wage in agriculture	n.i.	n.i.	0.097
<i>Environmental factors</i>			
Share of rural population	0.016	0.019	-0.176
R ²	0.79	0.84	0.72
Number of observations	223	300	176
Number of regions	10	13	9

n.i. Not included.

R². Coefficient of determination.

Note: Variables are expressed in logarithm—the marginal impact is the percentage impact of a 1 percent difference in the variable; region dummies are employed; the bold font indicates statistical significance at 10 percent and higher.

Source: Authors.

Appendix Table 4.1.3 Computation of the Healthcare Expenditure Need, 1997

Name	0-6 y.o. (%)	7-9 y.o. (%)	10-15 y.o. (%)	16-17 y.o. (%)	Pensioners (%)	Wage (%)	Budget coverage (%)	Service need (index)	Service cost (index)	Service responsibility (index)	Expenditure need (RUR)	Actual expenditure (RUR)
Samara City	-31	-30	-14	-15	-7	73	-95	0.92	1.10	0.72	296.7	343.2
Rostov on the Don	-22	-33	-15	-15	-10	11	-98	0.89	1.02	0.76	282.5	199.4
Volgograd City	-22	-28	-15	-18	11	57	-72	0.90	1.08	0.84	337.6	223.2
Tolyati City	-11	-19	-1	-11	-36	214	-95	0.93	1.29	0.72	354.8	352.8
Ulyanovsk City	-11	-10	6	-8	-27	23	48	0.97	1.03	0.49	203.3	202.5
Yaroslavl City	-26	-24	-18	-17	0	77	25	0.94	1.10	0.76	322.5	221.9
Novokuznetsk City	-16	-20	-8	-19	-11	142	8	0.93	1.19	0.86	396.2	374.5
Kemerovo City	-17	-23	-7	-2	-15	122	-4	0.92	1.16	0.86	382.7	462.9
Tula City	-32	-26	-15	-13	11	16	35	0.93	1.02	0.76	300.1	196.2
Lipetsk City	-15	-15	1	11	-8	79	-14	0.95	1.11	0.68	296.6	326.3
Ivanovo City	-27	-31	-23	-19	-7	17	53	0.91	1.02	0.71	273.9	197.0
Orel City	-22	-24	-14	-17	-14	9	46	0.93	1.01	0.66	256.4	207.9
Taganrog City	-19	-27	-15	-36	2	9	-98	0.91	1.01	0.76	288.4	221.0
Volzhskiy City	-16	-15	3	-2	-12	53	-44	0.95	1.07	0.84	353.8	225.8
Shakhty City	-14	-25	-8	-26	11	89	-99	0.90	1.12	0.76	315.9	189.9
Prokopyevsk City	-18	-23	-10	-12	3	111	62	0.92	1.15	0.86	376.5	478.1
Novocherkask City	-24	-36	-25	-35	-9	73	-99	0.88	1.10	0.76	304.0	176.4
Volgodonsk City	-13	-18	12	9	-37	24	-99	0.93	1.03	0.76	302.4	213.7
Syzran' City	-19	-23	-5	-28	3	65	-92	0.92	1.09	0.72	294.8	379.2
Maikop City	-19	-21	-1	-25	-12	-8	81	0.93	0.99	0.59	223.5	224.0
Belovo City	-9	-16	4	9	-7	153	-48	0.93	1.21	0.86	400.8	390.9
Leninsk City	-11	-16	-1	-14	0	108	16	0.94	1.15	0.86	383.8	406.7
Kamyshin City	-9	-5	-14	-15	-10	-4	-31	1.01	0.99	0.84	348.5	288.0
Novoshakhtinsk City	-18	-35	-23	-36	11	130	-99	0.87	1.17	0.76	319.9	264.7
Elets City	-19	-35	-28	-27	-2	2	8	0.88	1.00	0.68	247.9	298.0
Kiselevsk City	-10	-13	8	-8	3	111	-26	0.94	1.15	0.86	386.1	383.2
Novokuibyshevsk City	-23	-23	-12	-3	3	112	-74	0.93	1.15	0.72	316.8	380.9
Anzhero-Sudzhensk City	-16	-24	-2	-12	8	73	-40	0.90	1.10	0.86	353.2	286.8
Kineshma City	-19	-24	-11	-15	4	0	56	0.92	1.00	0.71	269.1	339.3
Min	-32	-36	-28	-36	-37	-8	-99	87%	0.99	0.49	203.3	176.4
Max	-9	-5	12	11	11	214	81	101%	1.29	0.86	400.8	478.1

Source: Authors.

5

The Assignment of Revenues at the Subnational Level

After the assignment of expenditure responsibilities among government levels, the manner in which revenue sources are assigned to the same levels is generally considered the second “pillar” or component of a system of intergovernmental fiscal relations.

This chapter considers the revenue assignment in the Russian Federation, with a particular emphasis on the impact of revenue assignments on regional-local relations. A key finding is that despite an initial decentralization of revenue sources to regional and local governments during the early transition, revenue assignments have become increasingly centralized since the mid-1990s. An important concern that arises from this centralized revenue assignment is the limited revenue autonomy at the subnational level, which in turn contributes to the reduced local accountability of subnational government officials.

Our discussion of subnational revenue assignments in the current chapter is organized as follows. In the next section we outline the general principles and practices followed in assigning revenue sources among different government levels. This is followed by a review of the current legislation that guides Russia’s revenue assignment with a detailed description of each tax provided in appendix 5.1. Next, we consider the role of tax administration in the Russian tax system, which is discussed in greater detail in appendix 5.2. Since subnational revenue autonomy is a relevant concern in the Russian Federation, we develop a taxonomy of revenue sources in the Russian tax system based on the degree of subnational autonomy provided by tax instruments. Using this taxonomy, we consider variations in revenue composition across different regions by the degree of fiscal discretion and cross-jurisdictional differences in revenue autonomy. Consistent with our policy focus, in conclusion we highlight key policy issues in subnational revenue assignment in Russia and discuss changes in tax assignments recently proposed by the Kozak Commission. Finally, we present options for reform.

General Principles and Practices

Musgrave and Musgrave’s (1976) “three-roles” classification of public finance not only is useful in considering the assignment of expenditure responsibilities, but can also be used to guide the assignment of revenue sources across different government levels. After all, different tax instruments have varying impacts as to the three functions of the public sector: macroeconomic stabilization, redistribution of income, and resource allocation. It has been argued by Oates (1972) that macroeconomic stabilization and income redistribution should for the most part be a central-level responsibility. Hence, tax instruments that can significantly affect macroeconomic stabilization and income redistribution should be assigned to the central government.

For instance, most economists would argue that progressive income taxes should generally be assigned to the central government level, owing to the taxes’ redistributive nature as well as to the economic stabilization that results from applying higher rates on growing income. However, the same reasoning does not apply to flat-rate surcharges on the national income tax. By the same token, in order to pool the cyclical fluctuations in revenues, revenue sources that are highly sensitive to general economic conditions should be assigned to the central government. Thus, subnational governments should turn to more stable tax bases such as consumption taxes and property taxes.

As subnational governments are mostly prescribed to engage in activities ensuring a more efficient allocation of public resources, they should be assigned revenue sources for which it is easier to establish a link with the benefits received by residents from local government spending (Bird 1999).

The most obvious example of a revenue source satisfying this “benefit principle” is charging for specific services provided by local governments (auto registration, issuing driver’s licenses, and so on) and for goods and services provided by public enterprises (utility charges, museum admission, and so on). Beyond voluntarily paid fees, specific benefit taxes are compulsory contributions to local governments that are nonetheless related in some way to benefits received by the taxpayer. For example, front footage levies relate the individual taxpayer’s contribution to benefits received from street improvements. Finally, general benefit taxes can be exemplified by charges levied on motor vehicles and motor fuels, whose revenues can be used for the construction and maintenance of roads and highways and thus benefiting road users as a class. Likewise, property taxes are often considered a good benefit tax to finance local collective public goods.

Public finance theory and practical policy wisdom provide three desirable characteristics for local government taxes. First, as noted already, good local revenue sources generally link revenue collections to locally provided benefits. Second, good local revenue sources have a tax base that is relatively evenly distributed across jurisdictions. And third, good local revenue sources generally have an immobile base.

At least conceptually, the benefit principle should be the primary guidance for the assignment of local government taxes. The other two features are subservient to some degree to the benefit principle in local taxation. For instance, if the benefit principle is preserved in assigning local taxes, the potential mobility of taxpayers would not necessarily bring about economic distortions, since taxpayers would be receiving local benefits in accordance with local taxes paid. Meanwhile, the potential mobility of taxpayers and tax bases can actually increase efficiency in the delivery of subnational government services by forcing local and regional officials to provide a balanced basket of subnational services and subnational taxes. However, if non-benefit taxes are applied to mobile bases, inefficiencies can arise from tax avoidance costs, as taxpayers could try to reduce their subnational tax liability by moving between subnational jurisdictions without affecting the benefits received from publicly provided goods and services. Equally distorting, and therefore to be avoided, are local taxes that can be “exported” to taxpayers in other jurisdictions.¹ Although clearly tempting and attractive to individual local governments, the assignment to local governments of taxes that can be exported lead to inefficient and irresponsible behavior of local governments at a national scale.²

Relating taxes to the benefits of public spending has the major advantage of helping increase the accountability of subnational governments to their own constituencies. In jurisdictions where the level of taxation exceeds the value of the benefits of public spending, high-income individuals and investors could threaten to leave or never to come or invest there. Clearly, the benefit principle is violated when taxes are “exported.”

Horizontal or political accountability of subnational officials requires the ability of subnational governments to affect at the margin the level of their revenues by choosing tax rates for some of the most important taxes assigned to them.³ Limited subnational taxing authority and dependence on the revenue decisions of the upper-level government, including decisions concerning revenue sharing and most other forms of transfers, undermines the accountability of subnational governments to their constituency, by offering an easy “scapegoat” for their poor performance (“We do not get enough resources from above”) and by generally weakening local taxpayer awareness of taxes and interest on the quality and level of local services delivered.

1. “Tax exporting” is a situation in which the burden of taxes imposed by one subnational government is borne by the residents of other jurisdictions who do not benefit from services provided by the government levying the tax.

2. Examples of taxes that are likely to be exported in large part include excises levied on the production of alcohol or production taxes on local mineral resources that are consumed or utilized outside the local jurisdiction levying the tax.

3. Other forms of local tax autonomy, such as the ability to change tax bases or the ability to introduce or eliminate some taxes, are generally less desirable since they lead to increased compliance and administration costs.

A number of recent studies (Ter-Minassian 1997; Ebel and Yilmaz 2002) do suggest that outcomes of decentralized spending depend on the form of financing used for these expenditures, with a crucial aspect being the extent of control that local governments can exercise over the sources of their revenue. Revenue autonomy is important for subnational governments for higher accountability of public officials and efficiency of expenditures, or for the ability to mobilize revenues and expand or contract the budgets at the margin; in addition, a healthful degree of revenue autonomy at all levels of government is the only certain way to address vertical fiscal imbalance.

Current Legislation and Practice in the Assignment of Revenue Sources

Revenue assignments in the Russian Federation have been evolving on an ongoing basis since the beginning of the transition. Table 5.1 provides a summary of major legislative and regulatory decisions affecting subnational revenue assignments.

Table 5.1 Summary of Major Legislation Affecting Subnational Revenue-Raising Powers

Period	Development
Dec. 1991	Introduction of the <i>Tax System Law</i> providing a closed list of taxes authorized for levying by different levels of government
Dec. 1993	Transformation of the subnational share of the CIT revenue into a regional surcharge of up to 22 percent
Dec. 1993	Increase of the upper limit of the <i>enterprise assets tax</i> rate to 2 percent
Dec. 1993	Introduction of the <i>federal transport tax</i> (1 percent of payroll) entirely decreed for allocation to regional budgets at the point of collection
Dec. 1993	Federal government gives subnational governments <i>carte blanche</i> to introduce their own taxes beyond those enumerated in the <i>Tax System Law</i>
1994	<i>De facto</i> unification of rates for sharing tax revenue between the federal and regional governments and introduction of formula-driven federal transfers
1997	Introduction of the requirement that regional governments should, on average, meet some minimum sharing rates for major taxes with local governments (Federal Law N 126-FZ of Sept. 10, 1997)
Nov. 1997	Abolishment of the federal <i>transport tax</i> and subnational governments' right to introduce their own taxes beyond those enumerated in the <i>Tax System Law</i>
July 1998	Legislation authorizing a <i>regional sales tax</i> of up to 5 percent, requiring the elimination of the <i>regional education tax</i> and 15 minor taxes at the local level (Federal Law No. 150-FZ of July 31, 1997)
July 1998	Legislation authorizing a regional tax of up to 20 percent on presumptive income that would replace existing federal and subnational taxes for businesses with largest compliance problems (Federal Law No. 1548-FZ of July 31, 1997)
2000	Enactment of the <i>Budget Code</i>
2000–01	Replacement of the subnational share in VAT collections with transfers earmarked for the fulfillment of federal mandates
2001	Replacement of a local tax of up to 1.5 percent of enterprise turnover with a local surcharge on the CIT of up to 5 percent, and lowering of the ceiling on the regional surcharge to 19 percent
2002	Replacement of the local surcharge on the CIT with an entitlement to 2 percent of the federal CIT collections and lowering of the ceiling on the regional surcharge to 14.5 percent
2002–03	Temporal centralization to the federal budget of all proceeds from the <i>forestry fee</i> and <i>charge on users of water resources</i>
2002	Switch in taxation of mineral resources from the subsoil fees to the <i>extraction tax</i> , so that revenue retention is eliminated for local governments and reduced for regional governments
July 2002	Legislation authorizing the <i>regional transport tax</i> (levied on vehicle owners based on engine power)
Jan. 2004	Elimination of the <i>regional sales tax</i>

Source: Authors.

The current federal legislation does not allow any level of government to introduce taxes beyond those enumerated in the *Law on the Basic Principles of Taxation* (gradually replaced as chapters of the new *Tax Code* are enacted). The list of permitted taxes is broken into three categories: “federal” revenue sources, “regional” revenue sources, and “local” revenue sources. The detailed description of major taxes is provided in appendix 5.1. As of January 2004, the list of “federal” taxes includes the value-added tax (VAT), excises, custom duties, the corporate income tax (CIT), taxes and royalties on natural resources extraction, the personal income tax (PIT), the social tax (levied on payroll), as well as other minor federal revenue sources.

Classification of a tax into a particular category does not always determine the level of government that receives the proceeds from this tax. For instance, although the succession and gift tax is classified as a federal tax, traditionally its full revenue yield has been allocated to the subnational budgets at the point of collection (that is, fully shared on a derivation basis). At the same time, the land tax, which is classified as a local tax, is shared among all three levels of government. For some taxes (designated as “assigned”), revenue sharing is fixed in legislation. For others (designated as “regulated”), revenue sharing is determined by the higher-level government as part of the annual budget process.

Parts of PIT collection and excises on goods (and until 2001, also the VAT) are shared by the federal government with regional governments based on the point of collection for these taxes. In turn, regional governments can reallocate a portion of those revenues to their local government. Similarly, royalties on natural-resources extraction are in part assigned to regional governments (and until year 2002, also to localities) at the point of collection. In addition, subnational governments can determine a surcharge on the federal CIT of up to 17 percent for regional governments and 2 percent for local governments.

More important, the current Russian nomenclature for taxes says nothing about the degree of tax autonomy associated with the taxes assigned to each government level. As discussed below, the fact that a tax is classified as local or regional does not generally mean that either local or regional governments have discretion to change the structure of the tax, or even to determine the tax rate. Likewise, simply because a tax by law is assigned to one level of government in Russia does not mean that subnational governments cannot be given control over the base or rate of the tax.

After elimination of the sales tax in January 2004, the current list of “regional” taxes includes the enterprise assets tax, the forestry tax, the transport tax (levied based on engine power), and the gambling tax. Proceeds from all regional taxes (except the forestry tax) can serve as “regulated” sources of revenue for local governments, so that regional governments can share those revenues with their localities (often at negotiated or differential rates) on a derivation or point-of-collection basis.

As of January 2004, the list of “local” revenue sources includes the individual property tax, land tax, the individual entrepreneur’s patent, and the advertisement tax (see table 5.2).

In summary, the current federal legislation rigidly assigns tax instruments to each level of government and does not allow regional governments to devolve part of their taxing powers to localities. Therefore, in Russia currently the only two ways for regional governments to pass on fiscal resources to their local governments are through the assignment of discretionary intergovernmental transfers (either in the form of “regulated” revenue sharing or subvention) or through long-term entitlements of local governments to a fixed portion of the yield from “regional” taxes (or the regional share of federal taxes). This means that in principle there is no cross-region variation in revenue assignment for local governments. Nevertheless, owing to varying economic structure and different tax bases, revenue disparities are likely to arise among different localities (especially between urban and rural local governments). As result, we can expect varying dependence of local governments on “regulated” revenue sharing and intergovernmental grants from regional governments.

Figure 5.1 reveals that the share of local governments in the consolidated government tax collections follows the same trend as the local share in the consolidated government expendi-

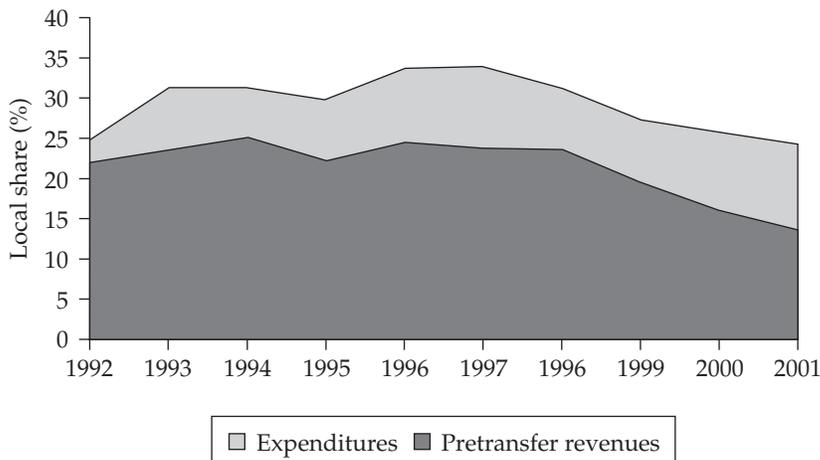
Table 5.2 Local Taxes and Fees

Tax	Tax is imposed by:	Tax base is set by:	Tax rate is set by:	Payments go to:
Individual property tax ^a	Federal government simultaneously for the entire RF	Federal government authorities	Regional and local	Local budget
Land tax ^a	Federal government simultaneously for the entire RF	Federal government	Regional and local authorities	Shared equally among regional and local levels
Registration fee on individuals engaged in entrepreneurial activities	Federal law simultaneously for the entire RF	Regional and local authorities	Regional and local authorities	Local budget
Advertisement tax	Local authorities	Federal government (cost of services)	Local authorities within federal limits 0–5%	Local budget

a. According to the General Part of the *Tax Code*, as soon as representative branches of regional governments put into effect a real estate tax in their jurisdictions, taxes referred to in the first two rows will cease to be levied.

RF. Russian Federation.

Source: Tax Code.

Figure 5.1 The Local Government's Share in Consolidated Government Revenues and Expenditure

Source: Author.

tures. Thus we observe a path toward more decentralization until 1996–97 and a steady decline in the local government share afterwards. However, it does appear that increases in the local governments' share in tax collections lagged behind those occurring on the expenditure side. By contrast, declines in the local share of revenue collections were leading those for expenditures both in time and scale. At the peak of decentralization in 1996, only 25 percent of consolidated government tax collections were directly allocated to local governments despite the fact that local governments accounted for 34 percent of consolidated government expenditures. By 2001, after five years of recentralization, these figures had dropped to 14 percent and 24 percent, respectively. Thus, tax revenues assigned to local governments, both from own-source and shared revenue sources, cover slightly more than half of their expenditures.

Tax Administration: Does Real Practice Fit the Law?

In the Russian Federation there is currently no autonomous tax administration by regional and local governments. Across the national territory, tax payments to all levels of government are collected by a local branch of the federal tax service. Then proceeds from taxes assigned exclusively to localities and regions are transferred to an appropriate budget while exclusively federal and shared taxes are transferred to a local branch of the federal treasury. Currently, it is the local office of the federal treasury that splits shared taxes among different budget levels according to legislated sharing rates. In this way the federal government keeps track of all collections in every locality to all levels of governments.

The current system of tax administration has its roots in the command economy and recent modernization efforts. Under the previous regime, taxes were administered by the finance department of the local Soviet (the local government council), which was subordinated to the finance department of the higher level of the administrative hierarchy. The bulk of budget revenue came from taxes (mostly on turnover, payroll, and profits) levied on and traditionally negotiated with state enterprises (Martinez-Vazquez and McNab 2000). In fact, taxes levied on State enterprises were substitutes to the remittance of their net profits to the state budget. As the owner of these enterprises, the State could ensure tax compliance in the course of annual audits covering all enterprises. Moreover, the State could routinely monitor bank accounts of enterprises through the State banking system.

During the Soviet period, taxes levied on individuals accounted for less than 10 percent of government budget revenues. Even for this category of taxpayers there was no need for sophisticated tax administration. Taxes on wages, retainers, and honoraria were nominally withheld at the source by the employers. The marginal category of individuals who were receiving income from non-State sources (for instance, the self-employed) could easily be identified and billed by finance departments.

The transition to a market economy required modernization of the tax administration. In January 1990, the Soviet Government decreed the establishment of the Chief State Tax Inspectorate within the Ministry of Finance as well as State Tax Inspectorates at lower levels of the administrative hierarchy. Two years later the State Tax Service (STS) was separated from the Ministry of Finance and directly subordinated to the Government and the President of Russia.

Parallel to the development of the tax administration apparatus, in 1992, a Presidential decree established the Chief Administration of Tax Investigation along with its 88 territorial divisions. In 1993, the Chief Administration was transformed into the Federal Service of Tax Police, an independent law enforcement agency, with some tax administration powers, among others. The Tax Police was charged with the investigation of tax fraud as well as other illegal activities such as illicit hard currency dealings.

The initial structure of the tax administration consisted of the State Tax Service Headquarters, 89 oblast-level State Tax Inspectorates (STIs), and almost 3,000 Territorial Tax Inspectorates (TTIs) at the city and rayon level. The STS Headquarters organizationally has been in a state of flux since its inception. Major reorganizations have been a common place and there has been a high rate of turnover of top-level management officials.

The total staff of the federal tax administration fluctuated from approximately 65,000 employees in 1990 to 197,000 in 1998 and down to 156,000 in 2000 (Tolkushkin 2001). The vertical structure of the federal tax administration appears to be extremely deconcentrated—even more so than the federal public administration as a whole. For example, in 1995 the STS Headquarters employed 710 persons while STIs and TTIs employed 161,000 persons. Interestingly, while the deconcentrated employment in STIs and TTIs accounts for more than one-third of the federal government's total public administration employment, the STS Headquarters staff accounts for less than 3 percent of the federal public administration core in Moscow. This highly deconcentrated structure of the federal Ministry of Taxation does not currently allow materializing all the economies of scale in the designing of procedures, processing of returns, and employment of specialized auditors. By

Box 5.1 *Do Subnational Governments Have Any Tax Administration Powers?*

It is argued that although territorial subdivisions of the federal Ministry of Taxation are not formally subordinated to subnational authorities, the former might harbor informal loyalties toward the latter. Thus, local authorities might have some control over how rigorously federal tax agents attempt to collect taxes in the local government's territory. The extent of this informal power and whether the federal governments have finally got a grip over their local agents continue to be discussed in the literature (OECD 2000).

Nevertheless, assuming that there is local control over tax administration, what implications does this have for the revenue-raising powers of local governments? There are arguments that actual taxing powers are determined to a greater extent by who actually collects the tax rather than by the formal tax assignment. Thus, the fact that in Russia tax administration is carried out by federal employees might imply that the actual taxing powers of local governments are smaller than could be assumed by looking at the legal tax assignment. However, the hypothesis of the localities' informal control over tax agents implies the reverse, namely that local governments in fact have some degree of control not only over their own revenue collections, but also over the federal and regional tax collections in their jurisdiction.

The next question is whether local governments' (informal) capacity to exercise assigned taxing powers is consistent throughout different regions. More important, to what extent is a possible variation in tax effort caused by regional authorities? The most likely causality line is through the hierarchical supervision of tax agents. Just as the tax agents might be influenced by the local authorities, their supervisors in the regional office of the federal tax agency might be (informally) controlled by the regional authorities. Thus, in fact we have a *triple* subordination of local tax agents: to the interests of the federal government, to the interests of their supervisors in the regional headquarters (who are in turn informally subordinated to the regional authorities), and to the interests of the local authorities. Everything seems to depend on which of the three is dominating. Moreover, there appear to be some tensions in the implied set of assumptions. For example, if the hierarchical supervision within the tax administration is strong enough to make local agents obey their regional supervisors, is it strong enough to make the regional supervisors obey their supervisors in Moscow?

We suspect that the more continuous contact between local authorities and tax administrators and the fact that the local authorities provide for important fringe benefits for tax administrators imply that the most important subordination is to local authorities, with subordination to regional masters only a distant second. Although Moscow may exert a lot of control over regional administrators, the pull is likely weakened as one goes down the hierarchy. However, it is likely that this situation is changing quite rapidly as the center continues the modernization of the tax administration apparatus and plans to introduce managerial innovations such as the frequent geographical rotation of local tax administrators.

contrast, STIs have been allowed and often encouraged to experiment with their own innovations and approaches (Firestone and Tax Administration Team 1998). The lack of coordination and logistical support from the Moscow headquarters traditionally has made federal tax agents in the field dependent on the support of local authorities. In turn this has contributed to blurring the delineation of powers among the levels of government and led to the *de facto* "dual subordination" of tax administrators to subnational and federal authorities (see box 5.1).

In 1999, the State Tax Service was transformed into the Ministry of Taxation. Since March 1999 the Ministry of Taxation has been in charge of maintaining the Unified State Register of Taxpayers, and it has been assigning Taxpayer Identification Numbers (TINs) to individuals since that date. Since January 2002 the Ministry of Taxation has been entrusted with the collection of social security contributions aggregated under the single social tax. Further consolidation took place in March 2003, when the Federal Tax Police was abolished and its functions were transferred to the Ministry of Interior.⁴

4. During the 10 years of its existence as an independent agency, the Federal Tax Police was deemed ineffective mainly because of the lack of coordination with the tax administration. The Tax Police had no institutional incentives to raise tax compliance. Instead, it pursued its own targets set by the federal government, such as the number of criminal cases launched and brought to court. While incorporation into the Ministry of Interior might improve coordination between the investigation of tax dodging and that of other economic crimes, it is not clear whether the move will have the proper impact on tax compliance.

Over the last decade efforts to modernize the operations of STS have been made with the support of the World Bank, IMF, United States Agency for International Development (USAID), and the European Union, among others. Reforms in the tax administration system have included moving to a functional organizational structure, introducing self-assessment systems for filing returns and paying taxes, increasing computerization, and improving taxpayer services. Although progress has been made, significant problems remain in Russia's tax administration system.

From the viewpoint of this study, there are significant problems with how the current structure of the tax administration and its reforms come to play with Russia's fiscal federalism. Some of these problems may be due to the lack of subnational tax administrations, or at least caused by the lack of focus of the federal tax administration on issues of subnational importance. Given the current way in which collection targets are set by the federal government, tax agents have no incentives to consider the subnational aspect of tax revenue allocation. For example, since 1998, enterprises have been required to pay taxes on the assets of their detached subdivisions (subdivisions of enterprises not registered as separate enterprises and not maintaining separate accounts) to the jurisdictions in which those subdivisions are located. However, it has been difficult to enforce this provision since neither the federal tax administration nor the enterprises have any incentives to split tax payments in this way.

Currently the federal government sets targets for the tax administration only in terms of the total collections by each tax. Therefore tax agents have little concern for which locality receives the local share of the payment from a particular taxpayer. Similarly, after the recent introduction of individual tax identification numbers, the tax administration has been empowered with the technical capacity to track the residence of persons for whom the income tax is withheld at source. However, the tax administration still has no incentive to allocate the funds from withheld taxes based on where taxpayers reside. A possible solution to this problem—the establishment of subnational tax administrations—is of limited use in the case of Russia, where tax policy is practically monopolized by the central government. According to the normative considerations and international experience discussed in appendix 5.2, in such an environment centralized tax administration brings significant gains from single procedures and forms designed at the central level and applied throughout the country uniformly and at little additional cost. This also reduces compliance costs for taxpayers. Furthermore, economies of scale can be achieved through return processing, record maintenance and retention, and by employing specialized technology, equipment, and specialized staff such as auditors. By contrast, in separated local tax administrations it is less feasible to organize staff by function rather than by tax thus undermining internal controls. Moreover, central administration facilitates rotation of personnel and eliminates close contact between auditors and taxpayers.

The only areas where independent subnational tax administration can be warranted in Russia are property taxes and user charges. While the administration of these revenue instruments does not present cost advantages associated with the size of government, this is precisely the area where local governments have some discretion on rates and bases and thus tailoring tax administration to local conditions can be beneficial. Moreover, the information needed to administer the property tax differs substantially from that which would be required, say, for other direct taxes, such as corporate or personal income taxes, or for indirect taxes, such as the VAT or sales taxes. Indeed, the focus of a property tax administration is to generate a register of property titles, as well as valuations that are reasonably up to date. Thus, if there were a single tax administration for all taxes, it is likely that property taxes would be collected separately from, for example, the VAT. Therefore there is less opportunity for savings from administering the property tax jointly with other taxes.

Taxonomy of Revenue Sources by Degree of Autonomy

As we noted at the beginning of this chapter, ensuring adequate revenue autonomy at the subnational level is a key challenge in Russia's transition toward a new federalism. However, it is not easy to quantify the degree of revenue autonomy local governments have. Depending on

the budget classification system and the definition of specific revenue items in a country, it is not always true that reporting the composition of local government revenue by type of tax does in itself reveal the degree of control the local governments have over these taxes. This type of problem clearly exists with the current Russian budget classification system on the revenue side. Nevertheless, it is important to know how much revenue autonomy exists at the local level today and whether the revenue performance in different regions of the Russian Federation can be traced back to degree of revenue autonomy exercised by their local governments.

One way to quantify the extent of local revenue autonomy in the Russian Federation is to explicitly classify subnational government revenue sources by the current degree of local discretion. Then we can use this taxonomy of revenue sources to track trends in local government revenue autonomy and to make cross-regional comparisons.

Conceptually, all revenue sources of subnational governments can be subdivided into the following categories ranked by the decreasing degree of control that subnational governments can exercise:⁵

- (a) Subnational authority sets the tax rate and tax base.
- (b) Subnational authority sets the tax rate only.
- (c) Subnational authority sets the tax base only.
- (d) Higher authority sets the rate and base of the subnational tax.
- (e) Tax-sharing arrangements fix the revenue-split in legislation for an extended period but this may unilaterally be changed by the higher authority.
- (f) Revenue-split is determined by the higher authority as part of the annual budget process.
- (g) Intergovernmental grants are determined by the higher authority as part of the annual budget process.

Subnational authorities have some control over the revenue sources in categories (a)–(d). In the remaining categories, local discretion is very limited or nonexistent. In table 5.3 we apply this taxonomy to the revenue items in the current Russian budget classification and report items that fall into OECD categories (a)–(e). In addition, we classify taxes by the discretion of local government to introduce a tax instrument, where Type 1 stands for such discretion to introduce a tax instrument, while Type 0 stands for the absence of this type of discretion.

As can be seen from table 5.3, most of the discretionary revenue items reported by local governments belong to either category (b) or category (e).⁶ Thus, for most of the revenue sources over which local governments have some control, the discretion is limited to setting the rates for the predetermined taxable bases. There is only one revenue item—the holiday resort fee—falling into category (a) so that local governments for this tax have discretion to choose the taxable base. Another curiosity is the revenue item “License for retail trade in alcohol,” which falls into category (d). For this revenue item, local governments can only decide whether to levy it or not, while the rate and base are fixed in federal legislation.

For a period between 1994 and 1997, as mentioned above, a presidential decree gave regional and local governments practically *carte blanche* to introduce their own taxes beyond those enumerated in federal legislation. Thus, for this period, we classify the “Other taxes and charges” item as belonging to category (a).

5. We follow the taxonomy suggested by the Working Party on Tax Policy Analysis and Tax Statistics of the OECD Committee on Fiscal Affairs. Note that the OECD taxonomy considers only those taxing powers that are actually exercised by at least one locality in the country. A different form of tax autonomy is the discretion to introduce any tax whatsoever. Although this type of autonomy to introduce any tax is quite uncommon, President Yeltsin granted it to subnational governments between 1993 and 1997. Below we explain how we capture this additional type of autonomy.

6. We base our classification on the *Law on the Basic Principles of Taxation* with subsequent amendments and gradual replacement with the provisions of the new *Tax Code*.

Table 5.3 *Classification of Local Government Revenue Sources, 2001*

<i>Budget classification code</i>	<i>Revenue item</i>	<i>Type</i>
1400501	Holiday resort fee	a, 1
1010110	Tax on profit (income) of enterprises and organizations decreed for local budgets	b, 0
1010111	Tax on profit of foreign legal entities decreed for local budgets	b, 0
1010112	Tax on profit of exchanges, broker agencies, credit and insurance companies; tax on profit from middleman operations and transactions decreed for local budgets	b, 0
1020433	Other license/registration fees decreed for local budgets	b, 1
1024000	Other taxes/fees on goods/services	b, 1
1040100	Individual property tax	b, 0
1050700	Land tax	b, 0
1400502	Earmarked charges on individuals, enterprises, institutions, and organizations for the upkeep of police force, landscaping, supporting educational institutions, and other purposes	b, 1
1400503	Advertisement tax	b, 1
1400504	Housing and utilities tax	b, 1
1400540	Other local taxes and charges	b, 1
2010000	Revenue from public property ownership and uses	b, 1
2060000	Administrative payments	b, 1
2070000	Fines and sanctions	b, 1
2080000	Revenues from foreign trade	b, 1
2090000	Other non-tax revenues	a, 1
1400505	License for retail trade in alcohol	d, 1
1020413	Alcohol production license fee decreed for local budgets	e, 0
1020700	Sales tax	e, 0
1040202	Enterprise Assets tax decreed for local budgets	e, 0
1040300	Estate and gift tax	e, 0
1050100	Levies on subsoil users	e, 0
1400100	State duty	e, 0

Note: See chapter text for an explanation of category (type) codes 0, 1 and (a)–(e).

Source: Authors.

As most items of the revenue classification are concentrated in just a few categories of the exhaustive taxonomy, for the convenience of our own analysis we merge the initial OECD categories into four broader categories: own-source revenue, assigned revenue, regulated revenue, and intergovernmental grants.

The category of “own-source” revenue encompasses all revenue sources whose yield can be affected at the margin by local governments—in other words, categories (a)–(d). This ability can stem from the local governments’ discretion to determine taxable bases, determine rates, introduce the tax, or any combination of these three. By exercising this form of fiscal discretion, local governments can adjust the level of public spending in response to local residents’ preferences. In addition to the ability to undertake fiscal adjustment in response to changes in resident demands, the presence of own-source revenues determines the ability of local governments to respond to changes in the costs of service delivery and economic fluctuations. To a significant extent, tax autonomy increases the ability of local governments to borrow in credit markets in order to undertake capital infrastructure projects. However, as pointed out earlier, the ultimate importance of own-source revenue is that it makes local officials more accountable to their constituency. This accountability is central to the good performance of a decentralized system of public finance.

The category of “assigned sources” of revenue refers to legislated long-term entitlements to (a share of) the yield from tax instruments over which local officials have no discretion.⁷ That is the category (e) in the OECD classification above. Although assigned revenue cannot be affected on the margin by local government decisions, it possesses two useful properties: it provides local governments with intertemporal predictability of revenue and ensures a revenue link to local government policy outcomes. The former property, intertemporal predictability of funds, helps local government to better budget and plan for future obligations and projects. In the Russian context, the latter property can be illustrated with the enterprise assets tax. Although local governments have no discretion over the tax base and rates, property development induced by local services provides local governments with fiscal rewards through their entitlement to half of the yield from this tax. This is an example of “fiscal” accountability” to the markets. Nevertheless, the assigned sources of revenue do not contribute to the accountability of local officials to their constituency. Even though local residents face the tax price of the local services funded with the local share of assigned-revenue sources, an efficient allocation does not occur because residents cannot respond to this tax price by choosing more or less of the local service.

The category of “regulated revenue” refers to tax-revenue sharing determined by the higher-level government as part of its annual budget process—category (f) in the OECD classification. Although, in terms of revenue autonomy, “regulated revenue” does not bring much improvement over discretionary intergovernmental grants, there are at least three aspects that can make regulated tax-revenue sharing different in the Russian fiscal system.⁸ First of all, tax-revenue sharing has more stability within the fiscal year, as the actual split is carried out by the Federal Treasury, which is guided only by legislated decisions. (Of course, regional legislatures can amend the sharing rates during the fiscal year; however, legislated decisions are less arbitrary than the process of disbursing discretionary grants.) Second, tax revenue shares are deposited into local government accounts at the end of each day while intergovernmental grants are disbursed at best on a monthly basis. Third, to the extent that past sharing rates serve as a benchmark for the current budget decisions, the rigidity of “regulated revenue” can bring about some positive qualities inherent in assigned revenue.

The final category in our pared-down taxonomy, intergovernmental grants, compiles all discretionary transfers to local governments whether formula-based or completely discretionary. Of course, these sources of revenue bring the least revenue autonomy to local governments.⁹

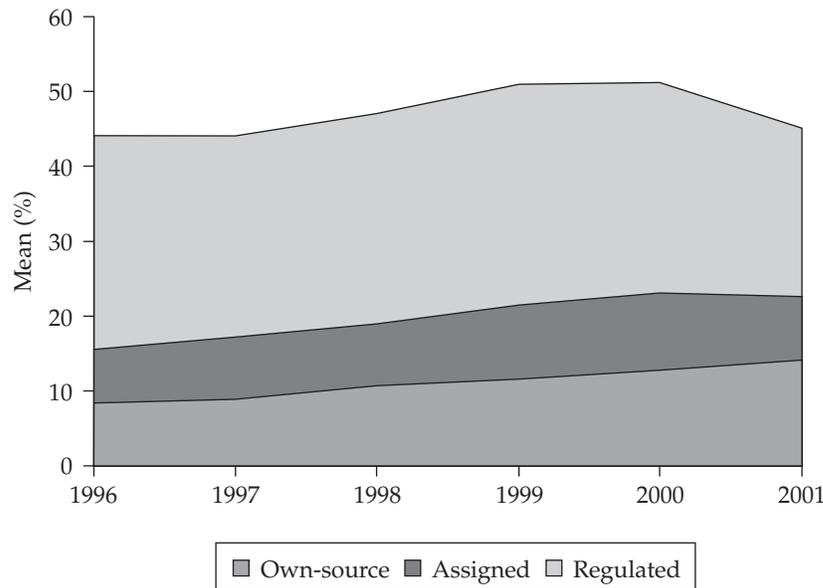
Distribution of Revenue Yield by the Degree of Fiscal Discretion

In the previous subsection we showed how local government revenue items are distributed across our taxonomy for measuring different degrees of revenue autonomy. But obviously each of the revenue items has a different yield and one tax instrument might generate more revenue than a dozen other taxes. Therefore it is important to quantify the degree of autonomy that exists in relation to the different yields of the revenue items. To this end we use budgetary data for individual local governments in different regions. In particular, we assess the mean shares of the four categories of the taxonomy of revenue autonomy (own-source revenue, assigned revenue, regulated revenue, and grants) during 1996–2001. Although the number of regions in the sample varies (from 24 in 1996 to 76 in 2001), this translates into a minimum of 700 localities

7. These statements would need to be qualified if the local authorities can affect the behavior and diligence of tax administration officials. The assumption here is that at the present time the dual subordination of tax officials in Russia does not present a problem any longer.

8. In general terms, revenue sharing is not more than a form of grant, distributed according to peculiar formulas. In the case of revenue sharing on a derivation basis, the formula is “in proportion to where the tax is collected.”

9. Strictly speaking, not all grants should be viewed as being equally negative in terms of revenue autonomy. For example, an equalization grant based on stable formulas for both funding and distribution of the funds can bring much more stability and predictability to local budgets than ad hoc specific grants decided every year.

Figure 5.2 *The Average Composition of Local Government Revenue (All Types)*

Source: Author.

for any particular year, which should produce credible descriptive statistics.¹⁰ We should further note that some of the local government revenue reported as “regulated” in this section might in fact have the nature of assigned-revenue sources. This would be the case if, for example, the regional legislation fixed local governments’ entitlement to a portion of the yield from “regulated” taxes for an extended period of time.

The evolution over time of “tax autonomy” at the local level in Russia is presented in figure 5.2 using the mean share of different revenue categories in total revenues of local governments in our sample. The figure clearly shows the increasing importance of own sources and assigned sources of revenue over time. The combined share of these two categories in local governments’ total revenue increased from 16 percent in 1996 to 23 percent in 2000. The share of own-source revenue alone increased from 8.4 percent in 1996 to 14 percent in 2001. Nevertheless, the importance of own-source revenue for local governments in Russia is still well below the levels observed in most OECD countries, as presented in table 5.4.

There are possible explanations for the increasing importance of own-source revenue. It is possible that over time local authorities have learned how to effectively use the available revenue-raising powers. However, a more likely explanation is the recentralization of expenditure responsibilities occurring since 1996 and associated with decreasing intergovernmental transfers (see figure 5.1). Local governments’ dependence on regional grants decreased from 56 percent in 1996 to 49 percent in 2000. Nevertheless, in 2001 the share of regional grants in local governments’ revenue increased again to 55 percent. This was a trickle-down effect from the federal government’s introduction of new earmarked grants in substitution for the former partial retention of VAT revenues by regional and local governments at the point of collection. Note that at the same time, the mean regional share of regulated revenues decreased from 28.2 percent in 2000 to 22.8 percent in 2001.

The discretion in setting the rates of about two dozen “local” taxes provides localities with some source of marginal revenue discretion. However, most of these taxes have low yield relative to their administration costs (for example, the tax on dog tags). Table 5.5 reveals that the bulk of

10. There is no guarantee, of course, that these unbalanced panel data are fully representative for all local governments in the Russian Federation during the observed years. Therefore we shall use caution not to extrapolate or forecast from the panel to all local governments in Russia.

Table 5.4 Revenue Autonomy of Local Governments in OECD Countries, 1995

	Percentage of own-source revenue	Composition of own-source revenue (percent)			
		Non-tax	Income and profit taxes	Property taxes	Consumption taxes
<i>Federal Countries</i>	53	28	12	10	3
Austria	50	42	0	3	3
Belgium	47	8	16	17	5
Germany	55	39	11	5	0
Mexico	22	22	0	0	0
Switzerland	82	33	42	7	0
United States	62	21	2	30	9
<i>Unitary Countries</i>	48	20	12	12	1
Czech Republic	25	20	0	5	0
Denmark	57	10	45	2	0
Finland	67	32	33	2	0
France	71	17	—	21	—
Hungary	29	23	0	1	5
Ireland	37	21	0	16	0
Italy	39	14	0	15	0
Netherlands	28	20	0	5	3
New Zealand	89	36	0	52	1
Norway	18	17	0	1	0
Poland	50	33	0	14	3
Spain	57	15	10	21	6
Sweden	82	12	70	0	0
United Kingdom	28	16	0	11	0

— . Not available.

Source: Derived from OECD (1999) and *Crédit local de France* and *Crédit communal de Belgique* (1997).

local governments' own revenue collections comes from just two tax instruments—the land tax and land lease; and the housing and utilities tax (of up to 1.5 percent on enterprise turnover), the latter of which was substituted with a local surcharge on the CIT in 2001. Moreover, for most of the local taxes for which localities have control over the rates, local governments have already set the tax rates at the maximum level allowed by federal legislation. Thus, in reality localities can only decrease (but not marginally increase) the size of their budgets through control over the tax rates. Furthermore, frequent changes in the ceiling for tax rates worsen the predictability of subnational revenue. Evidence of this is provided, for example, by the lowering of the ceiling on the local surcharge on CIT from 5 percent in 2001 to 2 percent in 2002.

The relative importance of particular sources of marginal revenue in local governments' total revenue is a product of both the productivity of authorized tax instruments and the political capacity to tap particular tax bases. Apparently, out of the two dozen "local" own taxes (over which local governments have some form of discretion), the most utilized are those on enterprise turnover, profits, and land use—all of them essentially business taxes. In contrast, the yield of the earmarked duties—which could be imposed on households like a poll-tax¹¹—has been decreasing since 1996. As of 2001, the average yield of the individual property tax remains at the negligible level of 0.4 percent of total revenue of local governments (although this is double its level in 1997).

11. The amount to be collected for local "earmarked charges" is set by the local government within federal limits of 3 percent of minimum wages for individuals and 3 percent of the minimum wage multiplied by the number of employees for enterprises. However, we are aware of only a few (unsuccessful) attempts to levy this charge on individuals, which would constitute a poll tax.

Table 5.5 *Average Structure of Local Government Revenue*

	1996	1997	1998	1999	2000	2001
<i>Own-source revenue</i>	8.38%	9.06%	10.83%	11.72%	12.95%	13.96%
	(0.67)	(0.62)	(0.63)	(0.66)	(0.70)	(0.84)
Profit tax	n.a.	n.a.	n.a.	n.a.	n.a.	6.86%
						(1.30)
Housing and utilities tax	3.69%	3.75%	4.76%	6.13%	7.88%	2.01%
	(0.90)	(0.87)	(0.89)	(0.94)	(0.97)	(1.14)
Land tax and lease	1.41%	1.95%	2.73%	3.18%	2.91%	2.56%
	(0.88)	(0.84)	(0.90)	(0.85)	(0.89)	(1.02)
Earmarked duties	0.97%	0.96%	0.41%	0.29%	0.19%	0.16%
	(0.98)	(0.94)	(2.99)	(1.44)	(0.90)	(1.77)
Non-tax	0.71%	0.63%	0.64%	1.20%	1.61%	1.66%
	(1.48)	(1.19)	(3.12)	(1.09)	(1.15)	(1.13)
<i>Assigned revenue</i>	7.30%	8.31%	8.27%	9.91%	10.25%	8.50%
	(1.05)	(1.11)	(0.99)	(0.89)	(0.95)	(1.13)
Sales tax	n.a.	n.a.	0.13%	2.74%	3.20%	2.68%
			(3.95)	(1.01)	(0.91)	(1.57)
Enterprise assets tax	5.86%	6.20%	6.51%	5.16%	4.31%	3.56%
	(1.09)	(1.14)	(0.97)	(1.12)	(1.13)	(1.29)
Subsoil fees	0.99%	1.72%	1.28%	1.67%	2.43%	1.95%
	(3.62)	(3.16)	(3.52)	(3.52)	(3.38)	(3.49)
<i>Regulated revenue</i>	28.57%	26.87%	28.08%	29.16%	28.16%	22.79%
	(0.56)	(0.58)	(0.59)	(0.61)	(0.63)	(0.67)
VAT	5.28%	5.09%	5.69%	4.78%	3.31%	—
	(1.00)	(1.02)	(0.96)	(0.94)	(1.07)	
PIT	13.02%	13.29%	13.55%	12.50%	13.73%	15.39%
	(0.57)	(0.54)	(0.61)	(0.59)	(0.60)	(0.65)
CIT	8.13%	5.73%	5.32%	7.79%	7.19%	2.64%
	(0.87)	(1.17)	(1.22)	(1.19)	(1.21)	(2.28)
Excises	0.55%	0.98%	1.11%	1.65%	1.15%	1.37%
	(4.84)	(3.76)	(3.60)	(3.57)	(4.33)	(4.20)
<i>Grants</i>	55.76%	55.71%	52.83%	49.22%	48.64%	54.75%
	(0.42)	(0.43)	(0.47)	(0.52)	(0.54)	(0.47)

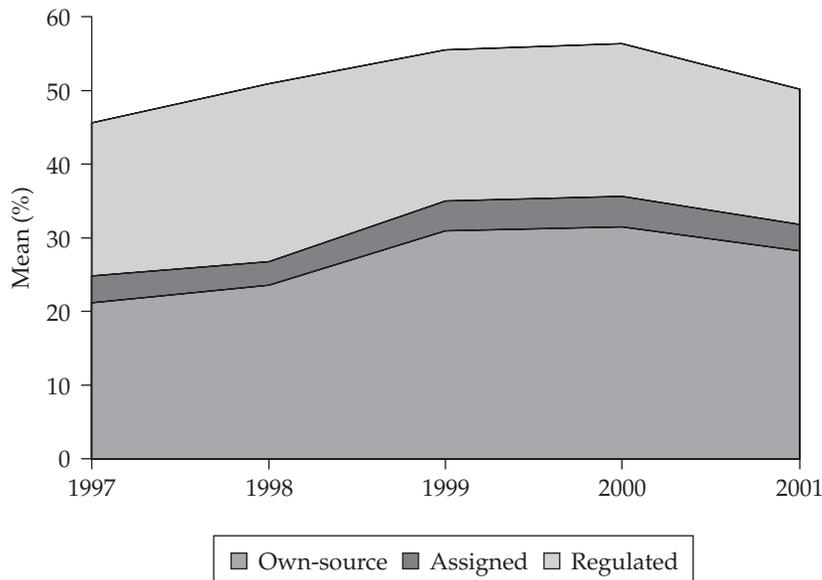
n.a. Not applicable.

Notes: Coefficients of variation are provided in parentheses. Averaging over all localities in our sample.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

There is one aspect of the more frequent use of the taxes applying to business that must be mentioned. To the extent that business taxes may be exported (paid by the residents of other localities), in effect this severs the link between paying taxes and receiving benefits from public spending, which is an essential element for sound local finance (Bird 1999). However, this may be exactly the reason why business taxes are more popular with officials and citizens alike than taxes more clearly paid by residents. Regardless of the final economic incidence of these taxes (that is, who really pays the tax in an economic sense), perceptions and political psychology play a role in the choice of tax instruments. Unlike more visible taxes on residential property and personal income, it is easy to assume (whether rightly or wrongly) that someone other than local residents pays business taxes.¹²

12. To the extent that businesses consume local public services it is also quite proper to have levies in the tax structure that applies to local business. In this instance, these levies are desirable even if the final burden is exported.

Figure 5.3 *The Average Composition of Regional Government Revenue*

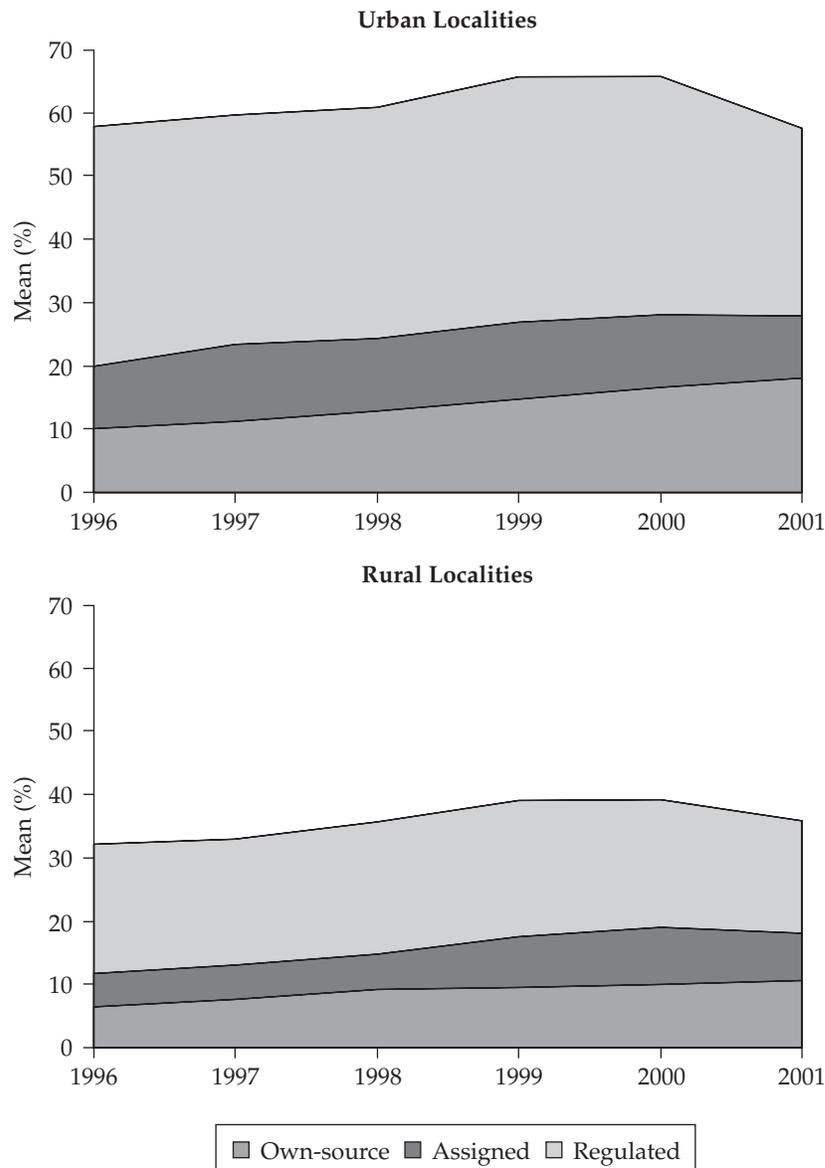
Source: Author.

The picture is a bit different for regional governments. Figure 5.3 reveals that regional governments actually enjoy a considerable degree of revenue autonomy. As of 2002, the regional governments' own-source revenue constituted almost 30 percent of total revenue, up from 21 percent in 1997. Moreover, some portion of the regional government total revenue is transferred to localities. Hence, relative to regional governments' own expenditure commitments, the proportion of own-source-revenue is even higher. Thus, many regional governments would seem to possess enough flexibility to undertake adjustments in response to changing costs of their programs or declining revenue bases. But note that as a result of this uneven distribution of fiscal discretion between the regional and local levels of governments, local residents might believe that the regional authorities are ultimately responsible for local government services. For the time being there is no survey or any other type of evidence that this is indeed the case. Naturally it would not be desirable for the efficiency of the decentralized system if local residents held regional authorities, and not local authorities, responsible for local service delivery.

Cross-Jurisdictional Differences in the Degree of Revenue Autonomy

Besides tracking the evolution of local government autonomy over time, our taxonomy allows us to make comparisons between localities. There is not likely to be any variation in the revenue-raising powers of local governments based on regional policies. After all, the current federal legislation virtually monopolizes tax policy and rigidly assigns tax instruments to each level of government—and it does not allow regional governments to devolve part of their taxing powers to localities. However, regional authorities can affect the mix of grants and regulated revenue received by different localities. These arrangements can also affect the yield of local governments' own revenue sources through the incentives to exercise the taxing powers available to them.

Revenue disparities between localities are more likely to arise from the structural differences in local economies. The same level of economic activity can be associated with different distributions of labor force and capital across different economic sectors such as agriculture, manufacturing, natural-resources extraction, or services. Different sectors of the economy may employ a different composition of productive factors, which in turn may be taxed at different rates. In particular, we can expect significant differences in revenue capacity between rural and

Figure 5.4 *The Average Tax-Revenue Composition of Urban and Rural Local Governments*

Source: Author.

urban jurisdictions. For example, federal legislation prescribes lower tax rates for taxing land in rural areas than in urban areas. In addition, until recently, federal legislation exempted agricultural business from many subnational taxes (for example, taxes on property and profits).¹³

Indeed, the comparison of mean shares for urban and local governments presented in figure 5.4 reveals drastic differences between the two types of local jurisdiction.¹⁴ The mean shares of all categories of discretionary revenue in urban jurisdictions exceed those prevailing in rural jurisdictions on average by a factor of 1.5. This disparity persists throughout all years, as both

13. Moreover, since 2002, regional governments have had the discretion to use a special single tax for agricultural producers in substitution for all current direct taxes. Originally the special tax was supposed to be levied on the land value but in 2003 the federal government proposed to change the tax base to "revenues less outlays."

14. As noted earlier, we classify a jurisdiction as "urban" if more than half of its population is urban.

types of jurisdictions follow the same trend. That is, for both the rural and urban localities, the mean share of their own and assigned sources of revenue has been increasing while dependence on regional grants has been decreasing (with some back-sliding observed in 2001).

After controlling for the structural differences in local economies, we can attempt to assess region-specific differences. In order to do that we compare the coefficients of variation (C.V.) in shares of different revenue categories *across all localities* with the coefficients of *between-region* variation in the same shares averaged for each region. Table 5.6 shows the composition of local resources by region. In addition, the table's descriptive statistics note not only the mean and

Table 5.6 Mean Share of Local Revenues in 74 regions, 2001 (percent)

<i>Oblast code</i>	<i>Subject of RF</i>	<i>Own-source</i>	<i>Assigned</i>	<i>Regulated</i>	<i>Grants</i>
1	Republic of Bashkortostan	12.7	11.1	12.1	64.1
2	Republic of Buriatia	7.7	7.3	10.7	74.3
3	Republic of Dagestan	2.7	0.8	4.4	92.1
4	Kabarda-Balkar Republic	7.2	2.8	33.4	56.7
5	Republic of Kalmykia	10.7	7.2	13.2	68.9
6	Republic of Karelia	11.8	9.3	31.6	47.4
7	Republic of Komi	12.9	11.7	27.2	48.2
8	Republic of Mari El	12.9	4.6	22.0	60.6
9	Republic of Mordovia	13.2	11.9	21.1	53.8
10	Republic of North Osetia	7.6	3.0	28.6	60.9
11	Republic of Tatarstan	19.9	11.4	15.3	53.5
12	Republic of Tyva	1.2	1.4	7.9	89.5
13	Udmurt Republic	11.0	19.1	22.4	47.6
14	Ingush Republic	2.7	2.2	14.9	80.2
15	Chuvash Republic	5.8	7.4	10.0	76.8
18	Krasnodar Krai	28.9	6.6	27.6	36.9
19	Krasnoyarsk Krai	11.1	4.8	15.5	68.7
20	Primorski Krai	11.8	6.5	18.2	63.5
22	Khabarovsk Krai	10.2	9.3	23.8	56.7
23	Amur Oblast	7.2	6.0	22.0	64.8
24	Arkhangelsk Oblast	13.6	3.3	32.4	50.6
25	Astrakhan Oblast	16.0	8.1	26.5	49.4
27	Bryansk Oblast	17.5	7.4	18.8	56.4
28	Vladimir Oblast	17.6	7.7	21.4	53.4
29	Volgograd Oblast	11.6	9.8	25.6	53.1
30	Vologda Oblast	17.9	5.3	30.5	46.3
31	Voronezh Oblast	21.0	6.9	25.0	47.2
32	Nizhny Novgorod Oblast	17.3	4.5	18.1	60.1
34	Irkutsk Oblast	10.6	8.2	43.5	37.7
35	Kaliningrad Oblast	10.9	5.1	32.3	51.6
36	Tver Oblast	9.0	5.9	26.8	58.3
37	Kaluga Oblast	10.9	4.1	32.7	52.4
38	Kamchatka Oblast	5.0	3.5	11.9	79.6
39	Kemerovo Oblast	10.2	5.0	23.3	61.5
40	Kirov Oblast	13.6	6.3	21.6	58.6
41	Kostroma Oblast	14.4	8.0	28.6	49.1
42	Samara Oblast	16.6	17.3	27.3	38.9
43	Kurgan Oblast	6.5	6.3	17.4	69.8
45	Leningrad Oblast	38.4	9.3	18.0	34.3
46	Lipetsk Oblast	17.4	10.7	25.9	46.0
47	Magadan Oblast	14.8	10.7	28.7	45.7
48	Moscow Oblast	27.9	12.4	28.7	31.0

(Continues)

Table 5.6 (cont.) Mean Share of Local Revenues in 74 regions, 2001 (percent)

<i>Oblast code</i>	<i>Subject of RF</i>	<i>Own-source</i>	<i>Assigned</i>	<i>Regulated</i>	<i>Grants</i>
49	Murmansk Oblast	9.7	7.8	37.2	45.3
50	Novgorod Oblast	9.2	5.5	36.6	48.8
51	Novosibirsk Oblast	12.3	9.2	21.8	56.7
52	Omsk Oblast	11.7	8.0	13.7	66.6
53	Orenburg Oblast	13.6	18.1	26.4	41.8
54	Oryol Oblast	8.1	4.2	29.5	58.2
55	Penza Oblast	8.6	3.2	17.7	70.5
56	Perm Oblast	15.1	13.5	24.1	47.3
57	Pskov Oblast	19.3	8.4	25.7	46.7
58	Rostov Oblast	17.4	3.1	21.4	58.2
59	Ryazan Oblast	13.4	7.1	27.2	52.3
60	Saratov Oblast	14.3	16.6	28.2	40.9
61	Sakhalin Oblast	15.3	8.9	27.6	48.2
62	Sverdlovsk Oblast	15.8	9.2	45.8	29.2
63	Smolensk Oblast	12.0	8.3	34.8	44.9
64	Tambov Oblast	11.6	6.2	8.4	73.8
65	Tomsk Oblast	13.5	14.3	13.8	58.5
66	Tula Oblast	7.9	4.5	20.9	66.7
67	Tumen Oblast	10.3	6.5	12.4	70.9
68	Ulianovsk Oblast	10.4	7.3	29.8	52.5
69	Chelyabinsk Oblast	17.1	12.0	22.1	48.9
70	Chita Oblast	11.6	8.5	12.6	67.3
72	City of St.-Petersburg	35.8	38.5	3.0	22.8
76	Republic of Adygeya	12.4	4.5	15.8	67.3
78	Jewish AO	6.6	1.8	10.6	81.1
79	Karachayevo-Circassian Republic	3.5	2.8	13.2	80.6
80	Republic of Khakasia	13.7	6.4	14.9	65.1
82	Komi-Perm AO	6.8	2.5	6.2	84.5
83	Koryak AO	8.3	6.8	17.3	67.7
87	Khanty-Mansi AO	20.1	26.4	37.7	15.7
89	Evenk AO	1.3	6.7	13.2	78.9
90	Yamal-Nenets AO	12.3	17.8	28.2	41.8
	Grand mean	12.8	8.3	22.2	56.7
	Median of means	11.9	7.2	22.0	56.5
	Total variation				
	out of which	28.0	18.9	47.3	137.2
	between regions	8.8	5.4	16.8	42.6
	within regions	19.2	13.5	30.4	94.6

AO. Autonomous oblast.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

median of the relative share of each revenue categories, but also the regional variations in the composition of local revenues. Furthermore, the total variation is further broken down into variation between regions and within regions, as explained in box 5.2. From the two bottom rows of table 5.6, we can see that the within-region variation in the shares of all categories of discretionary revenue is more than double that for the between-region variation.

As we can see in table 5.7, for urban jurisdictions the coefficients of both the total and between-region variations are smaller than the similar coefficients calculated for the combined sample. On the other hand, for rural jurisdictions all coefficients of variations are larger than corresponding coefficients calculated for the combined sample. In addition, in both subsamples, the between-region variation explains a larger part of the total variation than in the com-

Box 5.2 *Breaking Down Total Variation into Variations between Regions and within Regions*

In mathematical terms, the total variation computed for the entire sample comprises the sum of the squared deviations of local government shares s_i from the grand mean share :

$$\sum_i \sum_j (s_{ij} - \bar{s})^2$$

This total sum can be rearranged into two components:

$$\sum_i n_i (\bar{s}_i - \bar{s})^2 + \sum_i \sum_j (s_{ij} - \bar{s}_i)^2$$

Here, n_i denotes the number of localities in region i and \bar{s}_i stands for the region i 's mean. The first component, called "between-group variation," characterizes the average dispersion of regional means from the grand mean. The second, component, called "within-group variation," characterizes the average dispersion of localities' shares from the corresponding regional means.

Table 5.7 *Coefficients of Variation in Revenue Composition*

<i>Jurisdictions</i>	<i>All types</i>		<i>Urban</i>		<i>Rural</i>	
	<i>Total</i>	<i>Between regions</i>	<i>Total</i>	<i>Between regions</i>	<i>Total</i>	<i>Between regions</i>
Own-source	0.84	0.52	0.71	0.48	0.89	0.62
Assigned	1.12	0.70	0.89	0.57	1.36	0.83
Regulated	0.67	0.41	0.54	0.37	0.70	0.44

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

binned sample. Thus, by disregarding the differences in urban and rural economic structure, we seem to detect more variation within each region.

Table 5.7 further shows that in all the samples, the between-region variation seems to be greater for the share of assigned-revenue sources than the variations in the share of regulated tax revenues. Thus, similar to the within-region variation, the between-region variation seems to be due to environmental factors (such as natural-resource deposits) rather than regional governments' policies. However, note that the revenue category exhibiting the second-largest variance between regions corresponds to the "own-sources revenue" category, which is likely to be affected by fiscal incentives introduced with regional policies.

In summary, the purpose of this subsection was to measure the extent of revenue autonomy of local governments. We classified revenue receipts reported by local governments according to the degree of control that they can exercise over these revenue sources. We use this taxonomy to track the evolution of local governments' revenue autonomy throughout the 1996–2001 period. Although, as of 2001, the yield from local governments' own-source revenue, on average, accounts for only about 14 percent of local government total revenue, this proportion has increased since the mid-1990s. Moreover, there is a great variation in the extent of local government autonomy both between and within regions. Thus, the yield from local governments' own-source revenue exceeds one-third of the total revenue in every tenth locality in our 2001 sample. On average, the variation in revenue composition within regions is larger than the variation between regions. The relative importance of the within-region variation decreases if we look separately into a subsample of urban or rural localities. There seems to be more variation among rural governments than among urban cities both within one region and across regions.

Determinants of Cross-Regional Differences in the Revenue Autonomy of Local Governments

In the previous section we showed that the share of own-source and assigned revenue varies widely from locality to locality. A larger portion of this variation occurs within regions. However, after we control for the structure of the local economy, by examining urban and rural subsamples separately, we still have considerable variation both within and between regions. This raises the question of what determines the variations in local revenue collections and local revenue autonomy, both within and across regions. Unfortunately, we do not have sufficient local-level data to explore the determinants of the within-region variations in local revenue composition further. However, we have some regional-level data that can be helpful in explaining the between-region variation in the composition of local revenues and differences in the revenue autonomy of local governments across regions.

As we argued above, regional governments do not have formal authority to affect the taxing powers of local governments. However, regional authorities can affect the mix of grants and regulated revenue received by different localities. We are interested, therefore, in finding out what set of regional characteristics can potentially explain the decisions made by regional governments about the mix of grants and regulated revenue that impact local governments' revenue structure. In addition, we make an attempt to account for regional-level institutions that can influence how local governments utilize their taxing powers.

First, an important potential determinant of a region's local government revenue structure is the relative dependence of the consolidated regional government on federal transfers. As local governments account for a considerable portion of the regional-local expenditures, a share of federal grants is likely to be passed on to the local level and thus potentially reduce the relative importance of other sources of local government revenues.

A second variable that *a priori* may be expected to have an impact on local government revenue composition is the relative level of devolution of expenditure responsibilities. When local governments are assigned a *de facto* larger share of responsibility for subnational public services, other things being equal, local governments are forced to better utilize available revenue-raising powers. Thus, we use the local government's share in the regional-local expenditures to account for the level of devolved responsibilities and thus the need to mobilize revenue.

Third, the relative availability of natural resources should also affect the composition of local revenue structure. We employ the volumes of oil, gas, and coal extraction to capture the importance of the subsoil fee as an assigned source of revenue. Before the 2001 amendments to the *Law on Payment for Natural Resources*,¹⁵ 30 percent of the subsoil fee revenues were allocated to the municipality where extraction took place.

Fourth, we employ the gross regional product (GRP) to approximate the elasticity of tax instruments with respect to the level of economic activity. In addition, GRP should capture the income elasticity of regional redistribution policies.

Fifth, we also make an attempt to take into account the role of politics at the regional level. Identifying political trends and issues within the regions is difficult because of the lack of available data. During all transition years, there has been a big divide between reformist regions and traditional or conservative regions. To sort regional governments into those two camps we employ the number of votes cast for the Communist Party candidate in the 1996 presidential elections. We expect that the higher the percentage for this variable the more conservative the regional governments would be, because the candidate's supporters would lean toward the old-style centralized system of governance.

Finally, we attempt to account for the possible impact that differences in the vertical structure of local government may have on the composition of revenue. In particular, we employ a

15. Federal Law No. 2395-1 of Feb. 21, 1992, Article 42, decrees revenues from all hydrocarbon resources (oil, gas, and coal) to be split among federal, regional, and local governments, with these levels getting 40, 30, and 30 percent respectively.

Table 5.8 Determinants of Local Government Revenue Structure, 1999 (All Types)

	<i>Own-source</i>	<i>Assigned</i>	<i>Regulated</i>	<i>Grants</i>
Constant	0.2315 ^c (0.0538)	0.0345 (0.0430)	0.5005 ^c (0.1050)	0.2335 ^b (0.1217)
GRP	-0.0003 (0.0002)	0.0003 (0.0002)	0.0000 (0.0004)	0.0001 (0.0005)
Oil	-0.0095 ^c (0.0027)	0.0088 ^b (0.0033)	-0.0150 ^b (0.0066)	0.0157 ^a (0.0079)
Gas	0.0168 ^c (0.0040)	-0.0091 ^a (0.0049)	0.0226 (0.0152)	-0.0303 ^a (0.0153)
Coal	0.0003 (0.0003)	-0.0006 (0.0006)	-0.0005 (0.0015)	0.0008 (0.0014)
FEDAID	-0.1790 ^c (0.0269)	-0.0499 ^b (0.0237)	-0.2069 ^c (0.0707)	0.4358 ^c (0.0871)
LOCEXP	-0.0257 (0.0389)	0.0202 (0.0439)	-0.0942 (0.0932)	0.0997 (0.0988)
Communist	-0.0008 (0.0006)	0.0008 ^b (0.0004)	-0.0020 ^b (0.0010)	0.0020 ^a (0.0012)
LGTYPE	-0.0003 (0.0101)	0.0025 (0.0080)	-0.0143 (0.0195)	0.0121 (0.0258)
R ²	0.34	0.23	0.19	0.40

R². Coefficient of determination.

FEDAID. Federal grants.

LOCEXP. Decentralization of expenditures.

LGTYPE. Local government based on the traditional city/rayon level.

Note: Robust standard errors are provided in parentheses.

a. Statistically significant at 10 percent.

b. Statistically significant at 5 percent.

c. Statistically significant at 1 percent.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

dummy variable for regions where local government is based on the traditional city/rayon level as opposed to the settlement level (see chapter 3).

We employ these variables in an attempt to explain the local governments' revenue structure in 1999 for a sample of 58 regions for which data were available. The revenue structure for each region is modeled in four regressions, seeking to identify the determinants of the relative importance for each of the four broad local revenue categories in the overall composition of local revenues. The results of the statistical analysis are presented in tables 5.8 (all local governments), 5.9 (rural local governments only), and 5.10 (urban local governments only).

Surprisingly, the level of economic activity does not seem to have a consistent impact on the composition of local government revenue across all types of localities. However, in the subsample of urban localities it exhibits a positive impact on the share of assigned-source revenue. This is quite plausible and intuitive given that, after controlling for the subsoil resources (discussed immediately below), the assigned sources are dominated by sales and enterprise assets, both of which are likely to be positively correlated with economic activity and be concentrated in urban areas. For the rural subsample, the only significant impact of GRP is on the share of own-source revenue and this relation is negative. This is a surprising result but it may be explained in part by the fact that currently a large part of agricultural income escapes taxation altogether.

As expected, oil extraction has a positive impact on the share of revenue from assigned sources in all types of localities. Moreover, for rural localities oil extraction also boosts the share of received grants thus dampening own-source and regulated revenue.

Interestingly, the impact of gas extraction is quite different from that of oil and also differs by locality type. For urban localities, gas extraction has a negative impact on the share of assigned-source revenue.¹⁶ For rural localities, gas extraction has a positive impact on the share

16. This is likely to be explained by the fact that the recipient of Gazprom's property taxes is not determined based on the location of the company's properties but instead on the location of its headquarters in Moscow. Since 1998, enterprises have been required to pay taxes on the property of their detached subdivisions (not registered as separate enterprises and not maintaining separate accounts) to the jurisdictions in which those subdivisions are located. However, this requirement exempted enterprises engaged in transportation services. Because Gazprom does not separate its extraction and transportation activities, taxes for all of its properties are paid in Moscow.

Table 5.9 *Determinants of Local Rural Government Revenue Structure, 1999*

	<i>Own-source</i>	<i>Assigned</i>	<i>Regulated</i>	<i>Grants</i>
Constant	0.2194 ^c (0.0525)	0.0652 (0.0507)	0.3137 ^b (0.1567)	0.4017 ^a (0.2021)
GRP	-0.0006 ^c (0.0002)	-0.0002 (0.0002)	0.0012 (0.0010)	-0.0005 (0.0012)
Oil	-0.0097 ^b (0.0028)	0.0096 ^b (0.0044)	-0.0183 ^b (0.0074)	0.0184 ^a (0.0102)
Gas	0.0198 ^c (0.0031)	-0.0045 (0.0065)	0.0325 (0.0232)	-0.0478 ^b (0.0206)
Coal	0.0003 (0.0007)	-0.0004 (0.0013)	-0.0030 ^a (0.0017)	0.0031 (0.0019)
FEDAID	-0.1803 ^c (0.0259)	-0.0657 ^b (0.0265)	-0.0982 (0.1008)	0.3442 ^c (0.1255)
LOCEXP	-0.0310 (0.0371)	-0.0066 (0.0484)	-0.1766 ^a (0.0905)	0.2142 ^a (0.1178)
Communist	-0.0005 ^b (0.000)	0.0007 ^a (0.0004)	-0.0005 (0.0013)	0.0003 (0.0018)
LGTYPE	-0.0021 (0.0105)	0.0032 (0.0081)	0.0058 (0.0220)	-0.0069 (0.0312)
R ²	0.33	0.13	0.20	0.41

R². Coefficient of determination.

FEDAID. Federal grants.

LOCEXP. Decentralization of expenditures.

LGTYPE. Local government based on the traditional city/rayon level.

Note: Robust standard errors are provided in parentheses.

a. Statistically significant at 10 percent.

b. Statistically significant at 5 percent.

c. Statistically significant at 1 percent.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

Table 5.10 *Determinants of Local Urban Government Revenue Structure, 1999*

	<i>Own-source</i>	<i>Assigned</i>	<i>Regulated</i>	<i>Grants</i>
Constant	0.2217 ^c (0.0594)	0.0056 (0.0546)	0.5193 ^c (0.1129)	0.2533 ^b (0.1079)
GRP	-0.0001 (0.0002)	0.0005 ^b (0.0002)	-0.0007 (0.0004)	0.0003 (0.0005)
Oil	-0.0035 (0.0044)	0.0122 ^c (0.0033)	-0.0006 (0.0103)	-0.0081 (0.0094)
Gas	0.0011 (0.0091)	-0.0196 ^c (0.0037)	0.0287 ^b (0.0095)	-0.0103 (0.0128)
Coal	0.0002 (0.0010)	-0.0008 (0.0008)	-0.0025 ^b (0.0011)	0.0031 (0.0020)
FEDAID	-0.1825 ^c (0.0310)	-0.0563 ^b (0.0261)	-0.2234 ^c (0.0687)	0.4622 ^c (0.0772)
LOCEXP	-0.0594 (0.0494)	0.0041 (0.0551)	-0.0401 (0.1174)	0.0954 (0.1050)
Communist	0.0002 (0.0006)	0.0021 ^c (0.0005)	0.0001 (0.0012)	-0.0023 ^a (0.0012)
LGTYPE	-0.0001 (0.0114)	-0.0023 (0.0100)	-0.0403 ^b (0.0220)	0.0427 ^a (0.0245)
R ²	0.25	0.31	0.13	0.41

R². Coefficient of determination.

FEDAID. Federal grants.

LOCEXP. Decentralization of expenditures.

LGTYPE. Local government based on the traditional city/rayon level.

Note: Robust standard errors are provided in parentheses.

a. Statistically significant at 10 percent.

b. Statistically significant at 5 percent.

c. Statistically significant at 1 percent.

Source: Authors' calculation based on data from the Center for Fiscal Policy.

of own-source revenue and a negative impact on the share of grants. However, one should note that 88 percent of all natural gas extracted originates in just one region—the Yamal-Nenets Autonomous Area. Moreover, the extraction takes place mostly in urbanized areas. Hence, the estimate of the impact of gas extraction can be biased owing to the sensitivity of the statistical technique used here—ordinary least squares (OLS)—to outlier observations, in this case just the Yamal-Nenets Autonomous Area.

The extraction of coal does not seem to play any significant role in the relative yield of assigned sources of revenue. However, it exhibits a negative impact on the share of regulated revenue. This could very well be because most coalmines operate at loss and thus do not pay any profit tax.

As expected, dependence on regional transfers (included in the regression models as “federal aid”) is negatively correlated with the yield from all sources of pretransfer revenue. Expenditure decentralization (“local expenditures”) appears to have a statistically significant effect only for rural localities. In that subsample, devolved expenditures are negatively related to the share of regulated revenue and positively related to the share of grants. Thus, counter to our *a priori* argument, broader expenditure responsibilities do not make local governments better utilize their own sources of revenue, but rather they induce a significant increase in transfer dependence for rural local governments.¹⁷

The only consistent impact from a region’s opposition to reforms, as proxied with the Communist Party candidate’s performance, is on the share of revenue from assigned sources and this relation is positive. It is really hard to think how this causality could possibly work as assigned revenue by definition is outside of regional or local government discretion. In fact, here there might be some form of reverse causality, in the sense that regions with a lot of revenue disparities from uneven distribution of assigned tax bases might desire some redistribution.

In the sample pooling together both rural and urban governments, the Communist Party’s performance is also positively related to the share of grants and negatively to the share of regulated revenue. We speculate that these relationships are due to the close and positive correlation between the Communist Party’s performance and the relative importance of the rural population. This conjecture is supported by the finding that in the rural subsample, the effect of the Communist Party disappears. At the same time, in the urban subsample the performance of the Communist Party is significantly but negatively related to the share of grants.

As to the vertical structure of local government (“local government type”), the effect is detected only in the urban subsample in the form of a smaller share of regulated revenue and a larger share of grants. This is a quite interesting result as the regional choice for the structure of local government only has a direct impact on rural areas; regardless of which local government structure a region chooses, a large city would always be a single one-tier jurisdiction. The difference between a single-tier and two-tiered local government structures would be felt directly by villages and smaller subrayon cities. In a one-tiered structure, all subrayon entities would be pooled with other subrayon settlements into a single jurisdiction. In a two-tiered local government structure, these smaller cities would be separate jurisdictions but possibly subordinated to the deconcentrated arm of the regional government at the rayon level. One possible explanation of how the fragmentation of rural areas can affect large cities is that agglomerated rayons can exercise a stronger pressure on the regional governments to funnel regulated revenue away from large cities and redistribute it toward the rayons.

We have also tried in the regressions two other political variables that were eventually dropped because of their low explanatory power. First, we included the rating of investment risks by *Expert* magazine, hoping that it would measure political instability in the region, which also concerns the predictability of tax-revenue sharing. This variable turned out to be statistically insignificant in all regressions except the one for regulated revenue in urban areas. In that latter case the investment risk was negatively related to the share of regulated revenue. Thus, this result could be interpreted as indicating that in politically unstable regions, urban governments may tend to have less regulated revenue. Second, we included the share of the rural population to account for possible capture of regional politics by agricultural interest groups. However, the close and positive correlation between this variable and the Communist Party’s performance results in the statistical insignificance of these two variables when included jointly.

In summary, a number of trends emerge in our cross-regional analysis of revenue patterns, but only three regional characteristics have a consistent relation to the revenue composition of any

17. It is likely that the level of devolution of expenditure responsibilities and transfers are simultaneously related; therefore two-stage least squares or other more advanced estimation techniques may be required to disentangle these relationships.

type of local government. These variables are federal grants, oil extraction, and the performance of the Communist Party candidate in the 1996 presidential elections. Larger federal-regional grants reduce the relative importance of the shares of the pretransfer revenue sources; oil extraction boosts revenue from assigned sources; and the Communist Party candidate's performance also has a positive relation to the yield from assigned-revenue sources. Other regional characteristics show explanatory power only for either rural or urban governments, but not for both.

Main Issues in Revenue Assignments and Patterns

Revenue assignments at the regional and local levels have come a long way since the early years of the transition when the "regulation" of taxes was the most significant aspect of the yearly budget exercise, often bringing considerable uncertainty to the subnational budget processes. Successive tax reforms and the adoption of the *Tax Code* in 2001 have created much more certainty and rationality in revenue assignments at the subnational level. However, important problems remain. In this section we review the main issues and patterns in revenue assignments. In the following sections, we note current proposals to reform revenue assignments and propose several options for further reform.

Changing Revenue Assignments

Despite all the progress to date, the Russian Federation still lacks explicit and stable revenue assignments at the local level. Many regions continue to use a "regulating" approach to revenue assignments whereby local revenue shares from regulated taxes can be changed from year to year, which is intrinsically not different from the centralized, discretionary approach practiced in the former Soviet Union. In addition, during almost the entirety of the transition period, revenue assignments for local governments have been ever-changing.

Some Russian regions still see the "regulation" approach as the most suitable to deal with their realities, which include significant intraregional disparities in fiscal capacity. The two main advantages of the regulation approach are that it provides the regional governments with budget flexibility, and that it can allow for high levels of equalization. But, on the other hand, the regulation approach also imposes significant costs on local governments. First, regulation deprives local governments of any measure of revenue autonomy, as revenue autonomy and "regulation" of taxes are inherently contradictory concepts. Second, it diminishes certainty and predictability for local government budgets, thus reducing the efficiency of local governments and limiting their ability to develop strategic budget plans. Third, the practice can be used unfairly if some (richer or politically powerful) local governments are better able to bargain with the regional authorities. Fourth, the use of regulated revenues effectively claws back the fruits of local governments' revenue-raising effort and therefore creates perverse incentives for local governments to reduce tax collections, either by reducing their own taxes and fees or by relaxing the enforcement of shared revenue sources (see Alexeev and Kurlyandskaya 2003; Zhuravskaya 2000). And fifth, in some regions at least it appears that "regulation" is a source of ongoing political friction between regional and local governments.

Inadequate Level of Local Tax Autonomy

The current system of revenue assignments still provides local governments with little meaningful revenue autonomy. In reality, local governments do have their "own" revenue sources and they can set rates for some local taxes; however, these sources of revenue are (for now, at least) quite limited.

More important, the current system of incentives would seem to lead local governments in many regions to make little use of the limited revenue autonomy they have. Local governments

that may choose to use higher tax rates and fees are quite likely faced in the next budget period with lower revenue-sharing rates and grants from the regional government. One test of the impact of negative incentives on revenue mobilization is to analyze the extent to which local governments have exercised the tax autonomy they now have—in particular, how much use local governments have made of the local fees that they are currently allowed to introduce. We have seen that while the yield from the enterprise turnover tax has been rising, there has been little effort by local governments to utilize more visible local taxes such as the individual property tax or earmarked duties on residents.

The overall efficiency and equity of the system of intergovernmental fiscal relations in Russia could significantly improve with the formal separation of revenue sources between the regions and local governments. The choice for the federal authorities is whether to put these assignments into law directly, without giving regions any choice, or just to legislate mandatory principles or guidelines determining how regions need to organize revenue assignments within their jurisdictions. The former option is more common among unitary countries, bypassing entirely the role of intermediate levels of government in structuring revenue assignments for local governments. Consequently this approach may not be adequate for Russia.

Inadequate Tax-Sharing Mechanisms (CIT and PIT)

Revenue sharing on a derivation basis is the most important source of revenues at the subnational level after intergovernmental grants. This heavy reliance on revenue-sharing arrangements brings some potential problems. Revenue sharing (as opposed to the assignment of own-source revenue) to a large extent breaks the link between the benefits and costs of public spending. This may be a factor promoting general fiscal irresponsibility, as local governments have incentives to underutilize their own tax bases at the expense of national sharable revenues. Using the principle of “derivation” for revenue sharing introduces several additional problems. These are particularly acute for some taxes, such as the VAT and the corporate income tax (CIT). For the VAT, the division of collections among subnational jurisdictions is no longer a problem since this tax has been centralized and is now captured 100 percent by the federal government.

While the sharing of the CIT on a derivation basis presents several problems, there are several additional reasons why the CIT is less than an ideal tax to share with local governments. The CIT tends to be an unstable source of revenue, making it less suitable for local governments since local governments are (quite appropriately) not supposed to run budget deficits and have less of an ability to borrow to smooth out intertemporal fluctuations in revenue flows. The tax burdens from the CIT are also more likely to be exported to taxpayers in other communities, a feature that detracts from political accountability and responsible local governance. The tax base of the CIT is also less evenly distributed than are most other taxes, which tends to increase horizontal fiscal disparities at the local level.

In addition to these concerns, the sharing of the CIT with local governments is further complicated by how apportionment of the revenues is carried out in practice. Legally, the enterprise headquarters have to pay the CIT for their detached subdivisions to the jurisdictions in which those subdivisions are located. But the effective sharing of the CIT seems to be still based on the place of registration or the headquarters of the enterprise. This is relatively unfair since enterprise activities, and consequently the consumption of local public services, frequently take place in jurisdictions other than the one where the enterprise is registered. However, finding a remedy to this unfairness is not an easy task.

For quite a long time, the allocation of CIT revenues in the Russian Federation was frequently unrelated to where the actual economic activity took place. This led to a mismatch between the demand for public services and the availability of budget revenues to provide those services. In response to this problem the federal government introduced a formulaic approach to the apportionment of CIT revenues based on the geographical distribution of labor and capital inputs.

Two problems remain. The first is that taxpayers find the new formula complicated, and the federal tax administration does not have any strong incentive to enforce its use. The second issue is that the apportionment formula, when used, only mitigates and does not solve the problem of arbitrary revenue distribution from the CIT.

Although the personal income tax (PIT) is a much more appropriate tax to be shared with local governments, the apportionment of the PIT at the local level also presents problems. This is because the PIT withheld by enterprises on wage income is paid at the place of work rather than where taxpayers reside and where a larger share of the consumption of local public services arguably occurs.

Modernizing Local Fees and User Charges

A different aspect of revenue mobilization at the local level is the need to normalize and update user charges and fees charged for local services. Local agencies, including those involved in providing health and education services, have been making increased use of cost-recovery charges for the services they provide. The growth in the use of cost-recovery fees in fields such as health and education needs to be interpreted as a rational response from subnational authorities to cope with severe budget underfunding and limited revenue autonomy. In addition, revenues from fees may have the additional advantage of being protected from claw-back by the regional government. The unregulated practice of user charges may undermine accountability and may carry negative equity implications for the poor. On the other hand, this practice has contributed to increased revenue mobilization at the local level. These additional revenues have helped to fund, to varying degrees, the delivery of basic services that otherwise may have never been provided in some communities. In addition, the use of fees also has fostered cost-recovery practices in some areas of public services, which has the potential to improve the overall efficiency of the system of local government finance.

The rationalization of user charges will need to be accompanied by the reform of revenue assignments at the local level to increase tax autonomy. This rationalization needs to identify fees that are necessary and desirable and those that are not. In particular, rationalization of user fees should identify areas where full cost recovery is appropriate, so that fees that are charged to users of public services are commensurate with the actual cost of providing these services. For instance, local governments may need to increase their reliance on user charges to fully finance public utility services such as water and sewerage services, and in particular, public housing.

Deficient Tax Administration

Some problems exist with current arrangements that have a single (federal) tax administration. First, the allegiance of tax inspectors (who are federal employees) toward local authorities may still compromise tax collections and enforcement objectives. The evidence is that tax inspectors of the Federal Ministry of Taxation, even though appointed by the central authorities, still receive some fringe benefits from the local authorities, such as housing and other amenities. The danger is that, as experienced in other transition countries with similar institutional arrangements, local authorities may put pressure on tax inspectors to go easy with particular taxpayers, especially for taxes that are mostly assigned to the central government. Thus the question is whether Russia should have separate tax administrations at the federal and subnational levels. This issue is discussed in some detail in appendix 5.2.

Recently Proposed Changes in Tax Assignments

Recent reforms of the Russian tax system appear to be a natural response to the challenges that the federal government was facing in the aftermath of the economic crisis of August 1998. Recen-

tralization of tax revenues was carried out under the gloomy prospects of a debt payment peak in or about 2003. Despite predictions in the recent past of another financial crisis in 2003 similar to the crisis of 1998, the payments of the public debt have gone smoothly. The foreign debt of the federal government was reduced from 55.6 percent of GDP (US\$143.4 billion) in January 2001 to 34 percent of GDP (US\$122.1 billion) in January 2003. In addition, in 2003 the federal government quietly made debt payments totaling US\$15.7 billion, bringing the total debt to 24 percent of GDP.

In part, the avoidance of a new debt crisis was supported by the favorable world prices for oil during much of the postcrisis period. After 1999, the price for Russian oil remained above US\$20/barrel compared to the average of US\$12/barrel in 1998. The fear continues to be that if world oil prices were to fall again, the financial crisis of August 1998 could be repeated. Therefore the federal government has aimed the tax reforms at creating a favorable economic environment for the development of the manufacturing and service sectors of the economy. Overall, these reforms have tended to reduce the complexity of the tax system and lower the marginal tax rates, which also should reduce incentives for tax avoidance and evasion. While generally in line with sound fiscal practices, these tax reforms and changes in the assignments of tax revenues are causing substantial revenue losses for regional and local governments and have added very little to the revenue autonomy of these governments. In this sense, the recent tax reforms do not facilitate, and in fact may jeopardize, many of the potential benefits of a decentralized fiscal system.

Some of the most important recent tax reforms affecting subnational governments include the following:

- In 2001, the federal government substituted earmarked grants for the former partial retention of the VAT by regional and local governments at the point of collection.
- In 2002, a switch in the taxation of mineral resources from the subsoil fee to the extraction tax eliminated revenue retention by local governments of these taxes and reduced the share retained by regional governments.
- Subnational government turnover taxes, which used to provide a substantial level of own-source revenue for local and regional governments, were eliminated by January 2003.
- The regional sales tax was abolished in January 2004.

According to the proposed tax bills as spelled out in the reform proposals put forth by the Kozak Commission, the list of the regions' own taxes will include the enterprise assets tax, transport tax (levied according to engine power), and gambling tax. In addition, regional governments will be entitled to (a share of) the yield from the following federal taxes: 17 percent of the profit tax, 70 percent of the personal income tax, the presumptive income tax, the small business tax, excises on alcohol, excises on fuel and lubricants, the forestry tax, the water tax, charges on users of wildlife and water bioresources, the tax on the extraction of common natural resources, and 10 percent of the oil extraction tax.

The new list of local governments' own taxes, as spelled out in the Kozak plan, will include the individual property tax (assigned to settlement—that is, lower-level local—governments), the land tax, resort tax, the merchandise license fee, and the advertisement tax. In addition, local governments will be entitled to a share of the yield from the personal income tax (15 percent to districts, 15 percent to settlements), the presumptive income tax (30 percent to districts, 30 percent to settlements), the agricultural tax (30 percent to districts, 30 percent to settlements), and the small business tax (30 percent to districts, 30 percent to settlements).

As a result of the Kozak Commission proposal, the revenue autonomy of subnational governments will decrease from its limited levels in 2000 to miniscule levels after the proposed changes are phased in. The elimination of the local surcharge on the CIT (which was introduced in its day to compensate for the loss of the turnover tax revenue), earmarked charges, and other minor local taxes, will decrease the share of own-source revenue in the total revenue of an aver-

age locality from 14 percent in 2000 to 5 percent.¹⁸ The loss of local governments' entitlement to a share of the subsoil fee and the regional sales tax is expected to reduce the share of assigned revenue in the total revenue of an average locality from 10 percent in 2000 to 5 percent in 2004. Overall, the share of pretransfer subnational revenues is likely to be reduced from 51 percent to 35 percent of local governments' total revenue as the proposed changes are phased in.

Corresponding to the significant decline in local revenue sources, the vertical balance at the local government level might be roughly restored if simultaneously the responsibility for the financing (and possibly the provision) of education and healthcare is reassigned from the local to the regional level, as proposed by the Kozak Commission. In this case, in an average locality, as much as 90 percent of expenditures could be funded with pre-transfer revenue. However, the share of truly local own-source revenues would be still at a low level—about 10 percent. Thus, even if the stability of local government revenue might somewhat improve as a result of this reassignment, there would still be a big issue with revenue autonomy. The lack of discretion for local governments to increase revenue on the margin will continue to hamper their accountability to their constituencies and the overall efficiency of decentralized government. In addition, proponents of decentralization would argue that regional governments are likely to be less responsive and less accountable in the delivery of education and healthcare than are local governments.

While some view the losses in subnational revenue autonomy as the price paid for the recent macroeconomic achievements of the federal government, others do not see this tradeoff as inevitable or even positive. While the reforms eliminated a number of important (revenue-wise) tax instruments available to subnational governments, no new tax instruments have been provided to subnational governments in their place. And it is not that there are no good candidates for subnational taxes. In fact, there is some room for broadening the taxing powers of local governments by allowing surtaxes on PIT and excises, as discussed in the recommendations at the end of the chapter. (In addition, appendix 5.3 discusses options for various subnational “piggy-back” surcharges on federal taxes in Russia in greater detail). Overall, it appears that the federal government in general and the Kozak Commission in particular underestimate the importance of subnational governments' revenue autonomy for the fiscal well-being of the nation.

The proceedings accompanying the ongoing elimination of the regional sales tax bear witness to the general lack of awareness at the federal level of the importance of subnational revenue autonomy. The sales tax of up to 5 percent of cash purchases by individuals was introduced as a revenue-boosting measure after the 1998 crisis. The proceeds from this tax were decreed to be shared between the regional and local governments based on a 40:60 ratio. The new sales tax increased the share of regional governments' own-source revenue from 21 percent of total revenue in 1998 to 31 percent in 2000. At the same time the share of local governments' revenue from assigned sources increased from 8.3 to 9.9 percent. In November 2001, the State Duma approved the complete elimination of the regional sales tax by January 2004. In its proposal to Parliament, the federal government rightly quoted certain disadvantages of this tax, such as the increased compliance costs of a sales tax in the presence of the federal VAT, the possible cascading or snowballing effect of taxes on sales that are not completely exempted, and so on. But interestingly, the federal government never admitted or mentioned the advantages of the regional sales tax, including the importance of the sales tax for subnational revenue autonomy.¹⁹ On the contrary, federal authorities assured Parliament that the regions would be rightly compensated for the reduced

18. We arrive at these estimates based on an average structure of local government revenue in 2000. The crucial assumption here is that the yield from the remaining tax instruments does not change. Although the modernization of the property taxes eliminates some distortions, it is not expected to considerably boost revenue (see Timofeev 2004a). Additional revenue from raising the assets tax rate from 2 percent to 2.2 percent is likely to be offset by revenue losses from taking inventories and intangibles out of the tax base.

19. Appendix 5.4 provides a more comprehensive discussion of the advantages and disadvantages of having a federal VAT coexist with a regional sales tax in the Russian Federation.

revenue with transfers from the federal budget.²⁰ Moreover, these compensation grants—while being comparable in size to the federal equalization fund—are to be distributed on an *ad hoc* basis rather than on a formula basis, as is the case for the equalization fund.

This illustrates well the case of how the intergovernmental aspect is often overlooked in the formulation of fiscal policy by central governments. The Russian Federation is clearly no exception. More so, in Russia, tax policy authority is practically monopolized by the federal government. As a result, the legislated tax system basically cannot accommodate the various structures of local government established within Russian regions. The proposed bill by the Kozak Commission on local government attempts to bring some uniformity to the local government structure. However, it is not likely to make it much easier for the federal Ministry of Finance to control the finances of local governments from Moscow.

In fact, as we have seen, the federal Ministry of Finance does not directly deal with local governments and has to rely on the information received from regional authorities. At the same time, at least until 2003, local governments have carried out the bulk of subnational expenditures while the regional level mostly provided liaison between the federal and local authorities. Most of the subnational fiscal problems that the federal government has to solve by means of bailouts are concentrated at the local level (for instance, housing and utilities). The reassignment of the responsibility for financing some of the core subnational government functions (specifically, education and healthcare) from the local to the regional level might bring some improvement to the fiscal situation of many local governments and even to the effectiveness of federal policies. Indeed, the federal government can introduce conditionalities for the funds provided to regional governments, while it cannot do so directly with local governments. However, even if basic services such as healthcare and education could possibly be provided more efficiently at the regional level (which yet remains to be seen), the proposed reforms ultimately still fail to address the most important cause of vertical fiscal imbalance in the Russian Federation, which is the lack of adequate tax autonomy for regional and local governments. If anything, the proposed reforms would seem to worsen this serious problem.

Options for Reform

Several principles should be kept in mind when considering options for reform. The reform of revenue assignments should not be made in isolation from the rest of the system of intergovernmental fiscal relations. They need to be compatible with the rest of the decentralization system in terms of vertical balance and with the ongoing and planned reforms in expenditure assignments, reforms of the transfer system, and reforms of the budget process. For example, if the regional governments continue to insist on using negotiated minimum expenditure budgets and gap-filling transfers, these will be of no use in reforming revenue assignments. Of course, the reform also needs to be coordinated with tax policy and tax administration reform.

Increasing Revenue Autonomy

One of the first priorities in the reform of revenue assignments in Russia should be granting local governments a higher degree of tax autonomy. This will be the best way to address existing vertical fiscal imbalances and to increase the efficiency and accountability of local budgets. The least costly option to increase tax autonomy (in terms of compliance and administration costs and interference with the central government's ability to maintain macroeconomic stability) is to confine local tax autonomy to choosing the tax rate levels of a closed list of local taxes, keeping tax bases identical within and across regions.

20. According to Finance Minister Kudrin, "It is wiser to generate the compensation [for lost revenue from the sales tax] through subventions and transfers" (*Vedomosti*, June 6, 2003). Similarly, the head of the Budgetary Commission in Parliament stated that the "falling out" revenue of regional governments can be easily compensated by reshuffling revenue from federal taxes (*Vedomosti*, June 6, 2003).

There are not too many taxes that fulfill the desirable characteristics of local taxes, and at the same time provide adequate revenue for local governments; it is quite clear that not all taxes are created equal in terms of their suitability for local governments. In actual international experience, some form of personal income taxation and property taxes are the two most frequently used instruments to provide revenue autonomy to local governments (as highlighted by table 5.4). Some countries also rely, but much less compared to other revenue sources, on the taxation of consumption at the local level in the form of sales and excise taxes. Some basic choices for revenue autonomy at the subnational level in Russia are summarized below.

A PIGGYBACK PERSONAL INCOME TAX An obvious choice to provide local governments with some degree of tax autonomy is to develop a piggyback flat rate PIT. Larger local governments could be allowed to join with a piggyback rate of their own. A proportional flat rate PIT piggybacking the national (potentially progressive) PIT could be assigned to the local level. This tax should be (eventually) paid using the place of residence as basis, and the law should give local governments discretion to set rates between minimum and maximum bands legislated by the federal government. The flat rate PIT can be assigned to one single level of government (say, regions) or can be split in a prearranged legislated way among different subnational levels of government (say, among regions and lower-level local governments).

PROPERTY TAXES The best choice to develop local tax autonomy in Russia at the local level continues to be the introduction of a modern real estate (property) tax. However, from the viewpoint of revenue yield, the property tax will take some time to produce noticeable results. The key to the successful introduction of the new property tax will be the completion of the fiscal cadastre, the introduction of efficient market-based mass appraisal methods, and a fair and transparent administration of the tax, including efficient appeals procedures. Given the difficulties involved in completing the cadastre and introducing mass appraisal, the government may consider, for an interim period, assessment procedures that are less complex and sophisticated. Updating some of the presumptive methods now used for the land and property tax could yield significant improvements in tax revenues.²¹ Another way to enhance local tax autonomy associated with the real estate tax is to regulate the voluntary introduction by municipalities of “betterment” or improvement levies. These are surcharges to the property tax that municipalities may approve within their jurisdictions, as one-time or multiyear charges, for improvement directly benefiting certain homeowners, such as improvements in street lighting, sidewalks, and so on. These levies have become common in many states in the United States as well as a number of developing countries and in some cases represent a significant source of revenue.

SUBNATIONAL EXCISES AND A REGIONAL VAT Other less obvious options exist for enhancing revenue autonomy at the local level. These include allowing regions and large municipalities to introduce local surcharges on the excise taxes on alcohol, tobacco, and petroleum products. Where the size of jurisdictions is small and distances are short, differences in subnational excises raise the possibility of “smuggling.” To be effective, subnational excise taxes require the strict use of stamps at the factory or point of importation for the product to be sold in each jurisdiction. A different option for increasing subnational tax autonomy is the introduction of a regional (but not local) VAT with the same base and structure as the national VAT but with regional discretion to set the rate. However, there is little international experience with this type of tax and it requires a very well-developed tax administration and VAT enforcement apparatus (see appendix 5.4 for a more detailed discussion of the various options).

21. In summer 2003, the federal government proposed amendments to the *Tax Code* that essentially modernized existing property taxes. While this can indicate a lack of willingness ultimately to introduce the real estate tax, it is likely to boost revenue from the existing property taxes by updating the ten-years-old methodology of assessing the reconstruction value of structures owned by individuals.

BENEFIT LEVIES Local taxation should shift from relying predominantly on business taxes (turnover, payroll, nonresidential property) to relying on taxes more visible to the local constituency so that they could hold local government accountable. User charges and improvement levies on property are two examples discussed in this section. As mentioned earlier, surcharges on alcohol and tobacco excises can also provide a good link to the benefits of local government activities to the extent that the latter are responsible for healthcare. Similarly, the benefit principle accords with allowing local governments to piggyback on (destination-based) vehicle and fuel taxes to the extent that local governments engage in constructing and maintaining municipal roads.

Introducing Predictable Revenue Assignments at the Local Level

The lack of stable revenue assignments at the local level has not only contributed to the lack of local autonomy but also has deprived many local governments of the necessary budgetary certainty to plan their expenditure programs over time. Local governments' inability to predict their revenues continues to lead to poor planning, rushed decisions, and inefficient public spending at the local level. One of the most important steps in improving revenue assignments in Russia is to introduce stable revenue assignment (for a period of three years or more), including fixed sharing rates at the local level. The establishment of stable revenue assignments will require a careful exercise in balancing expected revenues with assigned expenditure responsibilities, while also requiring complementary measures. In particular, it will be necessary to neutralize other avenues regional governments can use to claw back lower-level governments' revenues. This latter will require establishing, among other things, a system of transfers based on formulas and objective and predictable criteria, and a culture of adherence to hard budget constraints for all dealings between regional and local governments.

Improving the Apportionment of Shared Taxes Adequately

Arguably the first best solution to the problem associated with the subnational apportionment of the CIT is simply to centralize all revenues from the CIT. This may not be politically feasible, so a possible second-best solution might be to share the revenues from the CIT proportionally across local governments where the business operates. An acceptable approach to the sharing of the CIT among the various jurisdictions in which the enterprise operates is to use an apportionment formula. Russia already has such a formula, whereby for taxation purposes, the total profits of a firm should be apportioned among jurisdictions in proportion to the subdivision's employment and assets—each factor weighted with one half. The problem is that the formula is not used all the time. But even if the profits apportionment among formal subdivisions is formally enforced, there is still an issue of affiliated enterprises that are formally independent. It is a common practice to minimize taxes by using transfer pricing among affiliated companies so that profits are concentrated in a company far from where the production takes place—often offshore.

A second issue in revenue sharing is that it would be desirable to adopt the principle of sharing PIT among local governments based on where workers reside. This is a more important issue in those areas where individuals may live and work in different local jurisdictions, or even in different regions (for example, Moscow City and Moscow Oblast).

Rationalizing the Current System of Fees

As noted earlier, there is a need to normalize or mainstream user charges at the local level. The share of non-tax revenue in Russia's local budgets is less than one-fifth of the lower limit observed in OECD countries (table 5.4). Partially this can be explained by the existence of extra-budgetary accounts where public entities accumulate nonbudgetary revenue. Apart from that, an increased reliance on user charges may be quite justified in an era of budget stringency, when local budgets are not providing sufficient funding for basic services. However, there are

two main concerns that need to be addressed. First, a system of user charges may exclude from service consumption the poorest groups of citizens. This would be counterproductive and could severely undermine the government objectives for income redistribution and fighting poverty. One solution is to use, where feasible, lower fees for the poor. A more general and preferable solution is to rely on income support programs for the poor that include at least enough income to pay for the higher services fees. Second, these user charges need to be brought out into the open, made official, and be regulated and controlled.

Streamlining Tax Administration

A final reform priority in the area of subnational taxation is for the government to continue the modernization and strengthening of the federal tax administration. In the medium term at least, local governments will likely need to continue to rely on the federal tax administration to collect their taxes. There are, however, three issues that need to be tended to.

First, the *de facto* subordination of tax administrators to local authorities can still create conflicts of interest now or in the future. It would be desirable to make it illegal for local authorities to provide any direct or indirect compensation (including housing or bonuses) to state tax inspectorate employees. For this measure to be effective, the federal government needs to adequately fund the budget needs of the state tax administration.

Second, federal government tax administrators may not see it as a priority now, or in the future, to collect local taxes and fees. The best way to address this problem is to arrange for a formal payment mechanism between the local governments and the central tax administration, by drawing incentive-compatible contracts between the two parties. For example, local governments may pay federal tax authorities a small percentage of local tax collections, with increased payments for increased collections.

Third, now is the time to start thinking about and preparing for the creation of regional and local tax administrations in the future. The objective should be to develop local tax administrations and to put them in charge of collections of local taxes and fees. The collection of piggyback taxes riding on national taxes could be left to the federal tax administration. Piggybacking is an incentive-compatible arrangement since the federal authorities would have an incentive to collect the national taxes anyway. The creation of a subnational tax administration could be done on a pilot basis. Eventually, where local governments are too small to operate their own tax administration at a minimum scale of efficiency, the option would be for those local governments to contract with a higher-level government for the administration and collection of their taxes. But again, these contracts would need to be incentive-compatible so that the higher-level tax administrators have an interest in and responsibility to collect local government taxes.

Appendix 5.1. Overview of the Current Tax System

Currently federal legislation prescribes a closed list of taxes that can be levied in the Russian Federation.²² This list comprises three groups of taxes: federal, regional, and local. When introducing taxes, regional and local governments must fit tax bases to the federal law and may levy rates only within federal limits. Certain taxes introduced by one level of government are shared on a derivation basis with another level of government (see appendix table 5.1.1). Revenue from some taxes are split among levels of government in proportions fixed in federal legislation while for other taxes the sharing rate is adjusted annually (such tax revenue is called “regulated”).

Since January 1999, the federal parliament has been enacting separate parts of the draft *Tax Code*. In addition to the general part of the *Code* dealing with tax administration issues and tax-

22. Articles 18–21 of the federal *Law on the Basic Principles of Taxation* (No. 2118-1; December 27, 1991).

Appendix Table 5.1.1 Legislated Sharing Rates of Major Taxes in the Russian Federation

	1999 Q1		1999 Q2–Q4		2000		2001		2002		2003		2004	
	Fed	Reg	Fed	Reg	Fed	Reg	Fed	Reg	Fed	Reg	Fed	Reg	Fed	Reg
VAT ^a	75	25	85	15	85	15	100	0	100	0	100	0	100	0
Profit tax	37	63	37	63	37	63	31	69 ^b	31	69 ^c	25	75	21	79
Personal income tax	14	86 ^d	14	86	16	84	1	99	0	100	0	100	0	100
Single social security tax	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	100	0	100	0	100	0	100	0
Energy excises	100	0	100	0	100	0	100	0	100	0	100	0	100	0
Excise on liquor	50	50	50	50	50	50	50	50	50	50	50 ^e	50	50	50
Excises on tobacco	0	100	0	100	0	100	0	100	0	100	100	0	100	0
Excises on wine, beer, and jewelry	0	100	0	100	0	100	0	100	0	100	0	100	0	100
Tax on property of legal entities	0	100	0	100	0	100	0	100	0	100	0	100	0	100
Land tax ^f	30	70	30	70	30	70	30	70	15	85	0	100	0	100
Tax on gas extraction ^g	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	80	20	80	20	100	0
Tax on oil extraction	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	80	20	80	20	86	14
Tax on extraction of common minerals	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0	100	0	100	0	100
Tax on extraction of other minerals	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	40	60	40	60	40	60
Tax on property of individuals	0	100	0	100	0	100	0	100	0	100	0	100	0	100

n.a. Not applicable.

a. VAT on imports, precious metals, and stones is assigned 100 percent at the federal level.

b. In 2001 a local enterprise property tax (EPT) rate of 5 percent was introduced; the regional share, therefore, also includes the local 5 percent.

c. During 2002–04 a local EPT rate was 2 percent, which was included in the regional share.

d. Regional share is taken as equal to average.

e. Since 2003, the federal government has been receiving its share of excises from manufacturers' sales to wholesalers while the regional share comes from wholesalers' sales to retailers.

f. Tax revenue from agricultural land was shared differently.

g. For regions including autonomous areas, the remittance rate is reduced by 5.5 percent in favor of the autonomous area.

Source: Modified and updated from World Bank (1996).

payer rights and obligations, those parts that develop federal taxes and the regional transport, gambling, and enterprise asset taxes are in force, as of January 2004. The parts of the *Code* that develop the structure of other regional and local taxes have yet to be discussed and approved by the federal parliament. In this appendix we describe major Russian taxes as of January 2004.

The Enterprise Profit Tax

During the early years of the transition, this tax was commonly used to promote or guide certain types of investment activities through either tax incentives and holidays or differential tax rates. More recently, as reflected in the *Tax Code*, Russia has followed the example of Western tax systems to provide a more uniform tax treatment of business activities across all sectors of the economy. The tax base of the enterprise income tax is calculated as the difference between taxable income and allowable expenses. Most companies use the accrual basis for this calculation. The calculation of the tax base of the enterprise profit tax has undergone profound transformations since 1991. Initially, Russia limited all deductions from enterprise revenues, including wages, capital depreciation, and interest, resulting in the tax being known in Russia as the enterprise income tax (rather than a corporate profit tax). The tax law currently allows the deduction of costs incurred in the generation of taxable income (for example, standard costs of doing business plus the cost of providing social services to workers). The combined tax rate is composed of the federal rate of 5 percent (6 percent in 2003), regional surtax of 13–17 percent (12–16 percent in 2003), and the local rate of 2 percent. Most taxpayers have to make monthly prepayments and file at least quarterly.

The Individual Income Tax

The current version of the individual income tax was introduced in January 2001. In designing this new tax, Russia elected not to adopt a global income tax that would have combined all taxable income into a single base. Rather, it elected to retain the features of a schedular income tax, which allows for different bases and rates depending on the source of income. This structure proved to be a judicious choice. The tax is levied on the worldwide income of residents and nonresidents. Personal allowances and certain hardship deductions are allowed, which are set as a multiple of the minimum wage. The base of the individual income tax in Russia includes all types of labor and employment income and income from certain other sources. The definition of employment income is broad, including fringe benefits, bonuses, allowances, and other forms of non-cash income. These forms of compensation are notoriously hard to tax at the individual level. Russia's personal income tax exempts income from pensions, some interest income, and capital gains from the sale of private property. The rate structure is flat at 13 percent of taxable income, with dividends subject to a rate of 6 percent, interest income on foreign currency deposits to 9 percent, and interest income in excess of the Central Bank refinancing rate and some material benefits to 35 percent.²³ Individual income tax is withheld at the source or paid directly by self-employed individuals.

Single Social Tax

This tax was introduced in 2001 to unify social security contributions by employers amounting to 28 percent of payrolls to the pension fund, 9 percent to the social insurance and medical funds, and 1.5 percent to the unemployment fund. Currently the cumulative rate structure ranges from 19.6 to 35.6 percent of payrolls.

23. Nonresidents' income is subject to a different rate schedule.

Value-Added Tax (VAT)

The value-added tax is the primary indirect tax levied in Russia. The Russian Federation introduced a VAT in January 1992 that was largely patterned after the one approved by the Supreme Soviet of the Soviet Union in December 1991. Two positive aspects of the Soviet-model VAT is that it had a single rate (albeit high at 28 percent) and a fairly broad base, covering most goods and services. However, the Soviet-model VAT presented many peculiarities and problems, some of which have been addressed in Russia and some of which remain in the current VAT. A major problem with the original Soviet-model VAT was that it denied credits for the VAT paid on capital inputs which amounted to 28 percent tax on investment. This practice destroyed the consumption basis of the VAT and introduced cascading elements in the tax. Currently, taxes on capital inputs may be recovered only after the capital project is completed (or sold before completion), as opposed to the practice of immediate recovery under the consumption-type VAT's utilized in most European countries.

Another important peculiarity of the Russian VAT is that the origin method applies not only domestically but also to trade with some Commonwealth of Independent States (CIS) countries. Exports of oil and gas to the CIS and all exports to Belarus are treated as domestic sales, so they are subject to tax, while imports from these countries are exempt from tax. In contrast, most countries with a VAT use the destination method for international transactions with third countries. Under that method, exports are zero-rated and imports are subject to tax. Taxpayers have a choice of accounting for tax liabilities for sales either on a cash or accrual basis. The number of exemptions and special treatments has significantly increased since the original adoption of the VAT. An extensive list of exemptions of both imported and domestically produced goods (which is considerably more extensive than that found in other countries) now exists. This practice results in a significant narrowing of the tax base and lower revenue yields. Apart from zero-rating for aforementioned goods, Russia has two VAT rates—a general rate of 18 percent and a lower rate of 10 percent for medicines, basic foods, children goods, and printed materials. Fraudulent claims for VAT refunds on sham export transactions present an increasing problem for administering the VAT in Russia.

Excise Taxes

Russia has introduced separate excise taxes on certain commodities (tobacco products, alcoholic beverages, and natural gas and petroleum products) and on “luxury goods” (large passenger cars). The excise tax applies to the sales by producer (in some cases by trader) as a flat fee per item. In harmony with the destination-based method used for the VAT for international transactions, Russia exempts excisable goods exported to non-CIS countries. Similarly, imports of excisable goods from non-CIS countries are subject to excise taxes.

Enterprise Assets Tax

Traditionally this regional tax has been applied to the annual average balance sheet value of assets of legal entities. However, after the adoption of chapter 30 of the new *Tax Code* in 2003, regional governments have had to reintroduce this tax to conform to the new federal framework. Thus inventories, intangibles, and capital outlays have to be excluded from the tax base. In addition, the federally imposed list of tax exemptions has been significantly shortened to include only enterprises of education, culture, healthcare, and sports.²⁴ Rates are to be set by regional legislatures and can vary by category of taxpayer (previously they varied by type of production for which the assets were used). However it is prohibited to set individual rates

24. Previously, assets used for agricultural production were exempted from taxation.

for single enterprises. The maximum rate is set at 2.2 percent.²⁵ Tax advances are paid quarterly after taxes are self-assessed by the taxpayer and returns are filed along with accounting reports with the tax inspectorate at the location of the legal entity itself and each of its detached subdivisions. The final payment is made within 10 days of filing the annual accounting report. Tax revenue is split equally between regional and local governments. Since 1998 enterprises have been required to pay taxes on the assets of their detached subdivisions (not registered as separate enterprises and not maintaining separate accounts) to the jurisdictions in which those subdivisions are located. However, it is difficult to enforce this provision since neither the federal administration nor the enterprises have any incentives to split tax payments in this way.

Land Tax and Land Lease Fee

In Russia, land use requires payment to the government. There are two forms of payment: the land tax and the land lease fee. The land lease fee is paid by lessees of State (that is, federal and regional) and municipal land. The amount and terms of payment are set separately in every lease. For State and municipal lands, the respective legislatures set basic rates by type of use and category of lessee. Land plots and shares of land plots in *ownership, possession, or use* of legal entities and individuals are subject to the land tax. Since it is classified as a local tax, the land tax is split between regional and local governments in proportions set by the federal government. The tax rate is set by local governments within a narrow range established by the federal and regional governments. Separate rates for agricultural and nonagricultural land are applied to the area of land plots. Tax rates are doubled for land plots not in use or used for purposes different from that intended. Regional governments can provide tax preferences to different categories of taxpayers and individual taxpayers on the portion of the tax liability due to the regional budget. At the same time, local governments are allowed to provide tax preferences on the portion of the tax liability due to their budgets but only with respect to individual taxpayers. In the case of individuals, the tax inspectorate at the location of the plot sends tax advice to the owner by August 1 of each year. Information on the area of plots comes from the *State Register of Land*. Information on their ownership comes from the *State Register of Property Rights*.²⁶ Legal entities make tax computations themselves and file their tax returns by July 1. Both legal entities and individuals pay the tax in two equal installments by September 15 and November 15. The payments are made to the local branch of the federal treasury, where it is then allocated to the two budgets. The new *Land Code* of 2001 establishes that taxation of land should be based on the cadastre value. The assessment of the cadastre value is commissioned by Russia's Federal Service of Land Cadastre (Roszemkadastr) and based on the mass-appraisal methodology developed by this federal agency. The federal government has been developing a new chapter of the *Tax Code* aiming at changing the taxable base from the acreage to the cadastre value in January 2005. However, as of 2003, 6 out of 89 regions have not developed a land cadastre. Just recently First Deputy Minister of Finance Shatalov announced that the land cadastre would be completed in 2004.

25. A review of 1998 budgets for 77 regions found that 55 levied the then-maximum 2 percent rate. Other rates levied included 1 percent (three regions), 1.2 percent (one region), 1.5 percent (two regions), and 1.8 percent (one region). Fourteen other regions levied rates that were differentiated by type of enterprise, with rates ranging from 0.05 to 2 percent.

26. All property rights acquired after the enactment of the federal *Law on the State Registration of Property Rights and Transactions* (No. 122-FZ; July 21, 1997) are subject to state registration. State registration of property rights is carried out by justice offices established by regional governments in accordance with federal legislation and under the supervision of the federal government. According to Malme and Kalinina (2001, p. 68), this registration process has been "... burdensome for new owners and costly for governments. . . . As a result, much property remains unregistered, including valuable new construction."

Appendix Table 5.1.2 *Federal Limits for Individual Property Tax Rates*

<i>Total value of property</i>	<i>Tax rates</i>
Below RUR300,000 (US\$10,000)	up to 0.1 percent
From RUR300,000 to RUR500,000	0.1–0.3 percent
Over RUR500,000 (US\$17,000)	0.3–2.0 percent

Source: Federal Law on Individual Property Tax (No. 2003-1; December 9, 1991) amended by Federal Laws No. 178-FZ of July 17, 1999 and No. 100-FZ of July 24, 2002.

Mineral Resources Extraction Tax

In 2002, the taxation of mineral resources was switched from the subsoil use fee to the extraction tax, which eliminated revenue retention for local governments and reduced it for regional governments. The objects of taxation are minerals extracted from a subsoil plot provided to the taxpayer under a user license. Minerals are broken down into those exempt from taxation, zero-rated, and taxed at non-zero rates. The tax base is self-assessed by the taxpayer by multiplying the unit value of extracted minerals by their volume. The unit value is periodically adjusted by the government based on the dynamics in world commodity prices. For oil extraction, the unit value is set at RUR347/ton during 2004–06. As of 2004 the tax rates applied to the self-assessed base range from 17.5 percent for gas condensate to 4 percent for coal.

Individual Property Tax

This tax, which is assigned to local governments, applies to structures (houses, apartments, cottages, garages, and the like) owned by individuals. The tax rate applies to the sum of the inventory values of all structures located in the jurisdiction and owned by one individual. If tax rates vary by type of structure, the rate is applied to the total value of all structures of one type and then liabilities are summed up for all types. The rates are set by the local government but with low federal maxima, which vary by the total value of structures (see appendix table 5.1.2). Regional governments can provide tax preferences to different categories of taxpayers as well as individual taxpayers while local governments are allowed to provide tax preferences only to individual taxpayers.

The inventory value of the structures is determined by a local Bureau of Technical Inventory (BTI), with the cost of reconstruction depreciated for the deterioration of the structure.²⁷ These bureaus are obliged to send information on the value and ownership of all registered structures to a local tax inspectorate by March 1 of each year. By August 1, tax inspectorates send tax advices to the owners. Payment should be made in two equal installments by September 15 and November 15. The late payment penalty is the overdue amount times 1/300 of the refinancing rate of the Central Bank for each day of delay. The proposed *Tax Code* chapter on the individual property tax reduces the tax rate ceiling to 0.1 percent of the inventory value. At the same time, the State Committee of the Russian Federation on Construction and Housing and Utilities (Gosstroj) updates its 10-year-old methodology for assessing the inventory value of structures owned by individuals, which brings it closer to the market value. The proposed bill significantly shortens the federal list of individuals exempt from this tax, and the federal government mandates abatements for the previously exempted category of individuals for structures valued under RUR500,000.

27. BTI is an authorized agent of the State Committee of the Russian Federation on Construction and Housing and Utilities (Gosstroj).

Appendix 5.2. Should Russia Have a Separate Tax Administration at the Local Level?

This appendix explores one of the fundamental questions in revenue assignments in the Russian Federation: whether subnational governments should collect their own taxes independently from the federal tax administration. Besides the obvious considerations of technical efficiency and administrative capacity, this issue also has significant political dimensions.

At the start of transition Russia had essentially an “upward flowing” tax administration system. This was the system inherited from the former Soviet Union. In that system, initially taxes were collected by the finance department of the local Soviet (council), which was subordinated to the finance department of the higher level of the administrative hierarchy. However, after each level of government gained political autonomy, tax administration officials found themselves under a “dual subordination” system, reporting to both the federal Ministry of Finance and the subnational governments exercising their new autonomy. Formally, this dual subordination was eliminated as the federal government established local offices of the Federal Treasury and Ministry of Taxation to serve federal needs in every locality. However, the federal government’s civil servants were recruited mostly from local finance departments, which often still paid for housing and other needs of tax officials. It is not surprising, therefore, that these officials were torn between competing loyalties. In the process, subnational governments were able to use this informal authority as a roundabout way to pursue their own tax policy. Thus, although the 1991 *Law on the Basic Principles of Taxation* initially decreed the revenue from the VAT to be entirely allocated to the federal budget, secure proceeds from this tax immediately became subject to sharing because of the subnational leverage on tax administration. More recently the federal government has been exercising more control over its tax agents. However, the problem of mutual mistrust among the levels of government remains.

Whether tax administration is centralized, allocated to a level other than the central level, or spread over various levels of government is likely to depend as likely on technical considerations as on political realities, but there are efficiency implications associated with different tax administration structures. Overall, there are both pros and cons of having an independent subnational tax administration. This appendix focuses on the technical issues related to the modernization of tax administration and the strengthening of fiscal federalism in Russia.

The Pros and the Cons of Separate Tax Administrations

There are four basic conceptual models for tax administration in a multitier government system: (a) central government tax administration only, with provision of revenue-sharing and transfers; (b) central government tax administration only, with assignment of different taxes to different levels of government; (c) multilevel administration, with revenue sharing and transfers; and (d) each level of government administering the taxes assigned to it.²⁸ In practice, these four models are not mutually exclusive, so some countries may not fit easily into one category. For example, Russia’s central government tax administration combines revenue sharing and tax assignment. In addition, the central government tax administration can have varying operational structures (for instance, by type of tax or by function) and varying degrees of deconcentration of authority to its territorial branches. Nevertheless, these conceptual models allow for a more focused evaluation of the tax administration issues.

The efficiency argument for the decentralization of tax administration is essentially the same as for decentralization of the public sector in general: having superior knowledge of local conditions, tailoring procedures to local conditions, constituting a laboratory for innovations. The counterarguments are also the same: inability to utilize the economies of scale, neglect of vertical and horizontal externalities, and lack of qualified staff.

28. See Vehorn and Ahmad (1997).

The primary drawback of independent subnational tax administration is the costs. Depending on the local government's billing and collection level, it can be expensive to have staff dedicated to collections only. Local governments can achieve some cost saving by integrating tax collections with the collection of user charges (including utilities and municipal housing rental charges). In contrast, a centralized tax administration is more likely to utilize economies of scale in return processing, record maintenance, and retention, and also be more easily able to employ technical specialists (for example, auditors). However, problems arise when the central tax administration has to cope with varying definitions of tax bases and different rates of tax across jurisdictions.

Under one organizational structure for all tax administrations, taxes are collected throughout the country following the same processes and procedures. This, among other outcomes, makes it easier for the tax administration to set up a simple computer system to monitor the collection of various taxes. Taxpayers' records with respect to each type of tax can be consolidated into one database or master file with a current tax account for each taxpayer. In this system stop-filers can be detected quickly, and delinquent accounts can be monitored in an efficient fashion. Enforcement, especially in its audit component, is simplified if the tax inspector has information, from the taxpayer master file, on every tax for which a given taxpayer is liable. Moreover, under a centralized tax administration there is no need to establish exchange-of-information agreements among tax administrations of different government levels.

However, centralized administration of subnational taxes can affect the behavior of tax officials, who may concentrate their efforts on collecting central government taxes at the expense of those allocated to subnational jurisdictions. For example, allocation of employees (such as highly skilled auditors) can be uneven if revenues from one tax were primarily distributed to local governments while revenues from another tax were primarily kept by the central government.

The subnational governments' advantages of having independent tax administration primarily concern control and accountability. The local finance office can directly supervise the collections staff and the procedures, and can institute control and audit mechanisms to ensure the honesty, efficiency, and accuracy of the collections office. Also, the subnational governments can oversee the customer's contact with the taxpayer or ratepayer in order to ensure that customers are treated politely and fairly. In addition, deposits can be made into the local government's bank accounts each day. Prompt depositing of funds is important for safeguarding and making the money available for cash management purposes.

Access of local governments to revenue and taxpayer accounts is instrumental for effective delinquency control (for example, the publication of a list of delinquent customers in local newspapers, termination of local government services, and so on). In some transitional countries with centralized tax administration the problem remains that municipalities do not receive adequate information from the national tax service. Tax legislation often forbids the tax administrator from disclosing any information to third parties, including municipalities, despite the fact that the corresponding tax revenues are allocated to local budgets.²⁹

Independent subnational tax administrations can serve as multiple laboratories for the modernization of tax administration; they deal with many small taxpayers, and they may select taxpayers for audit based on risk evaluation rather than checking every taxpayer at regular intervals. For example, in the United States, many states are using modern collection methods not currently used by the Internal Revenue Service (IRS) at the federal level. Some states have employed private collection companies to collect unpaid taxes and have allowed delinquent taxpayers to use credit cards in paying their taxes due.³⁰

The efficiency of tax administration is not only a matter of administrative costs but also one of compliance costs incurred by taxpayers. Taxpayer convenience is especially important for

29. In some counties in Estonia, the above-mentioned problem has been cleared by mutual agreements, as the national tax service in turn requires local governments to pass information on property objects for the administration of the (national) land tax (Högye and others 2000, p. 202).

30. See the United States General Accounting Office (1994).

taxes and tax systems based on self-reporting and voluntary compliance. Thus local governments might be better positioned to provide taxpayer service. However, the increasing use of the Internet and call centers makes this aspect less relevant. While giving local government officials increased autonomy, independent subnational tax administration can burden taxpayers with higher compliance costs if each jurisdiction issues different forms, regulations, and procedures. Compliance costs can be especially high for companies that have several branches or subsidiaries operating in different jurisdictional areas.³¹

The structure of tax administration concerns not only governance of the public sector but also the economy. Improvements to tax administration are instrumental for market reforms. Ten years of transition in Central and Eastern European countries show that economic recovery comes with the creation from scratch of small and medium-size businesses (World Bank 2002c). At the same time, large industrial enterprises at best produce the same level of value added, accounting for less than 40 percent of GDP. In a market economy, therefore, tax administration should be able to deal with large enterprises as well as many small taxpayers. These requirements are completely different from those associated with the type of tax administration inherited from the command economy.

An innovative approach to subnational tax administration can be helpful in establishing a broad taxpayer contact and reach tax bases beyond the large industry thus spreading the fiscal burden over all sectors of economy. However, it is not hard to imagine different scenarios, where for example arbitrary and harsh enforcement practices by the subnational tax administration (mostly against competitors of businesses close to the local elite) would upset federal government standards of fair treatment of taxpayers.

Administrative uniformity can simplify compliance costs and also increase the fairness of the tax system, in terms of the interpretation of the law and geographical equality in the selection of accounts for audit. It can be well argued that under centralized administration, taxpayers are more likely to receive uniform treatment in assessments, audits, penalties, and appeals, irrespective of where they reside. A more fair system is one that can bring about higher compliance rates by taxpayers. Thus multiple subnational tax administrations, other things being the same, may lead to lower compliance. But not all other things remain the same. When countries select multilevel or subnational tax administrations, authority to collect specific taxes should be given to the level of government that is able to function with the lowest collection and enforcement costs (Rubinfeld 1983). For example, it can be argued that individual and corporate income taxes and social security taxes can be collected more cheaply by a centralized tax administration. The same can be argued about collection and enforcement costs for the VAT, customs duties, and natural resource taxes. Subnational governments are generally more efficient in collecting other taxes, such as property taxes, certain excise taxes, and user charges.

One reason is that the information needed to administer the property tax differs substantially from that required to administer the corporate income tax or the VAT—and more important, it is of local nature. Thus, local tax administrations can be more efficient in registering property titles and updating valuations. However, uniform assessments may be easier to achieve where the assessment function is centralized. One study, for example, found that the use of county rather than local assessors resulted in more uniform residential assessments in U.S. jurisdictions. Another study suggests that, to the extent that there are economies of scale in the assessment function, these are more likely to be achieved at the central (state) government level. Of course, one way to achieve economies of scale while maintaining local assessment might be by contracting out the assessment function (Bird and Slack 2002).

The utilization of local information permits the effective local administration of licenses and other presumptive taxes on small businesses and informal sector activities. Such sectors are of

31. Vehorn and Ahmad (1997) report that it is not unusual for national corporations in the United States to fill out over 15,000 sales tax returns each year.

great importance in developing and transition economies and are relatively hard to tax. Countries with local tax administrations may also have an advantage in effectively using presumptive taxes since local governments can be more knowledgeable of the presumptive bases.

It should be noted that under multilevel tax administration, duplication of effort among levels can easily occur, and effectiveness can suffer if the various tax administrations do not put in place requirements for adequate coordination. Hence, coordination among various tax departments becomes crucial, especially in the area of audits.

In summary, the vertical structure of tax administration should be related to the assignment of taxing powers among different levels of government. If tax policy is monopolized by the central government, the gains from centralized tax administration are likely to be considerable. Single procedures and forms designed at the central level can be applied throughout the country at little additional cost. However, the important issue with a centralized approach is whether there are proper incentives for the central government officials to exercise rigor in collecting taxes assigned to subnational governments. Some incentive compatibility can be achieved through revenue sharing with the central government or withholding a percentage of local tax collections to pay the central government tax agency.

It must be remembered that even in cases in which subnational governments are given discretion to apply their own rate to the tax base used by the central government, as in a piggy-back subnational income tax, centralized administration still has its advantages. Administrative and tax-compliance economies can be achieved when local tax returns consist of only some extra lines on the national return or a supplement to it. However, the administrative economies disappear when the central tax administration has to cope with varying definitions of the tax base used by different jurisdictions. Inefficiencies in centralized tax administration can also result from centralized administration of taxes requiring knowledge of local information (for example, on property taxes or presumptive income taxes).

What Is the International Experience with Subnational Tax Administration?

The international experience presents many examples of independent subnational tax administration. Moreover, the list of such countries is very diverse—including the United States, the United Kingdom, Switzerland, Japan, India, Mexico, China, Germany, Hungary, Estonia, Albania, and others. The involvement of subnational governments in tax administration is to some extent related to their taxing powers, but this is not always the case. In particular, countries where subnational governments have legal powers to choose bases for their own taxes also tend to have independent tax administration (examples include the United States, Switzerland, and Japan). Federations such as the United States, Switzerland, and Brazil are examples of countries with multilevel tax administrations and tax assignments (with some control over revenue, administration, and legislation). Moreover, often the federal government and provinces have tax policy and administration powers over the same tax bases.

Among federal countries, Canada seems to do a good job of balancing the sovereignty of provinces with the economic gains from centralized administration of overlapping taxes. The federal government administers provincial individual and corporate income taxes and some provincial sales taxes are linked to the national VAT. Thus, with the exception of Quebec, personal income taxes are levied by both the federal and provincial governments, but collected by the federal government. Taxpayers file one unified (federal and provincial) tax return and calculate their provincial tax as a percentage of their federal tax (tax on tax). The percentage may vary depending on the taxpayer's province, because each province has the authority to set its own tax rate. If a taxpayer has moved during the year, the province of residence on December 31 receives the tax revenue for the whole year. The federal government pays an estimated amount to the provinces during the year and calculates a final reconciliation after all tax returns have been received.

Switzerland's case is quite different. Revenue is shared upward, and the cantons are responsible for the collection of most taxes. Thus the tax administrations of the cantons collect income

tax for both the canton and the confederation. While taxpayers only need to file one tax return covering both taxes, the definition of taxable income will normally differ between the canton and the confederation because each canton is free to determine levels of exemptions or deductions. So the taxpayer must make two calculations of taxable income, and the tax administration must verify that both calculations are correct. Also, tax auditors must be careful not to confuse the tax rules for cantons with the tax rules for the confederation. While the encouragement of diversity is the accepted practice in Switzerland, it does complicate tax administration.

In Germany, tax legislation is a federal matter, but tax administration is a state (*Länder*) responsibility. Customs control and tariff and duty supervision, however, are entirely federal concerns. Moreover, individual states can participate in tax legislation through the *Bundesrat* (Council of States). The *Bundesrat* members are directly appointed by state governments and can block decisions of the federal legislative house (*Bundestag*).³² In principle, tax administration is implemented on a uniform basis, following the federal law. Federal tax officials, in cooperation with state tax officials, determine tax administration processes and procedures. On a periodic basis, German tax administrators meet to harmonize implementation of the tax laws. The German *Länder* also have information-sharing agreements regarding businesses that pay taxes in more than one state.

Despite the successful experience of Switzerland and Germany with upward revenue sharing, this system can create perverse incentives to maintain revenue in the jurisdiction by a variety of means. For example, local tax administrators can reduce taxes going to the center by favoring local enterprises with low assessments and when it is permitted, by providing tax holidays and exemptions. This fundamental conflict of interest is what led the Chinese Government to introduce a separate tax administration at the local level in 1994, with the charge of collecting all locally assigned taxes, while the central tax administration was charged with collecting national taxes and also all shared taxes (see Bahl 1999).

It should be noticed that centralized tax administration with downward revenue sharing is not free from perverse incentives either. India is one of a number of countries with multilevel government tax administration, where tax revenue is shared downwards. In India, localities autonomously administer taxes on property and entertainment (movie releases, concerts, and the like), as well as the profession tax. However, central government tax officials collect all individual income taxes on nonagricultural income and then return 85 percent of revenue collected to the states. Even though the tax administration is given a collection target in the budget, the incentive to allocate sufficient enforcement resources and perform high-quality audits for this tax is lower than for other taxes such as the profit tax or customs duties, where revenue is retained by the central government.

Brazil illustrates the case of high administrative and compliance costs stemming from uncoordinated tax policies by different levels of government. In particular, there is an inevitable duplication of effort between the federal and state governments in the taxation of similar bases under the two VAT-type taxes. Taxpayers must have two taxpayer identification numbers and file two sets of tax returns. The tax administration must also follow two different sets of procedures for invoices of goods shipped between states and goods shipped within a state. Since there is no central statistical database to record sales and purchases for the state VAT, it is impossible to conduct a systematic cross-checking of information. Thus, state auditors are placed in a relatively weak position by a lack of important information.³³

Centralized tax administration is, not surprisingly, much more common in unitary countries than in federal countries. The tax administrations of Italy, France, Portugal, many countries in French-speaking West Africa, many formerly socialist countries, and certain Latin American countries fit into this model of centralized tax administration. In France, local taxes are determined by local authorities (within limits) but collected by the national tax office and redistributed. Also, in unitary countries it is more of a rule that overlapping taxes are administered by one level of government. Thus, in those Northern European countries (Finland, Sweden, Ice-

32. See Högye and others (2000, p. 64).

33. See Vehorn and Ahmad (1997) for further discussion of Brazil's system.

land, and Belgium) where the share of subnational taxation is one of the largest in the world, all taxes are administered by the central government. Moreover, recent reforms in Sweden have resulted in the centralization of certain functions within the national tax administration while staff was moved from the county tax offices to localities to enhance taxpayer contact. In Sweden, all income tax is paid to the central government. During the fiscal year, subnational governments receive their revenues based on a forecast. The final tax amount is settled in the following year when total revenue is assessed.

Denmark presents a mixed case, where the local income tax is assessed by the local government but collected by the central government. The local income tax is deducted at source by employers and paid by the self-employed in monthly installments at post offices. The same tax forms are used nationally and the same officials collect both national and local income taxes. Assessment is carried out by local committees that are appointed by municipalities but remain autonomous in their decisions. The municipal tax department acts as secretariat and calculates personal allowances and liabilities. Assessment is required to strictly follow national regulations. The central government commissioners of taxes exercise supreme authority for assessments, working through the Danish State Inspectorate. There are 65 boards of commissioners of taxes, and each of their five to seven members appointed by the Ministry of Internal Revenue, Customs, and Duties. They hear appeals on the assessment decisions of local tax committees. Taxes are collected centrally and then distributed to counties and municipalities according to taxpayers' places of residence and locally set rates.

Even among unitary countries there are examples of local governments collecting taxes shared with or piggybacked on by the higher-level government. In the United Kingdom, the uniform business rate (business property tax) is determined by the central government, but the revenue is collected by local governments. Each district council retains a sum up to the entitlement determined by the central government. The balance is remitted to a national pool redistributed by the Secretary of State for the Environment to equalize the resources of local authorities in accordance with their need as assessed by the Secretary of State's formula. In Japan, municipalities collect the prefectural local inhabitants' tax (per capita plus a percentage of income) along with their own and transfer it to the prefecture. For the municipal fixed assets tax, assessment is made by municipal tax officers in accordance with the national standard methodology.

In transition countries, the taxing powers of local governments are rather limited. However, it is often the case that local authorities autonomously collect fees and minor taxes that they are allowed to levy.³⁴ Thus, in the Czech Republic, local authorities only levy local fees. In Albania, localities autonomously administer taxes on property, small businesses, general sales, and local fees (for example, for parking). In Estonia, municipal and town governments independently collect local taxes in their territory. However, with the permission of the central office of the Tax Board, the municipal or town council and the local tax office may sign a contract whereby local income taxes can be collected by local offices of the national tax administration.

In Hungary, taxes are collected at two levels, central and local. Hungary's local governments are entitled to introduce local business taxes, communal taxes on entrepreneurs, communal taxes on individuals, and taxes on property (buildings), undeveloped plots of land, and tourism. Relations between central and local government tax administrations are not coordinated and there are difficulties in information flows between the two administrations.

What Are the Lessons for the Russian Federation?

The normative considerations and international experience suggest that central tax administration is a superior choice for the administration of subnational taxes that have a uniformly set base across all jurisdictions. In Russia, this would apply at the present time to corporate and personal income taxes as well as the road fund tax. This list would also have applied to the regional sales tax had it become harmonized with the federal VAT, instead of being eliminated. A central-

34. See Högye and others (2000, p. 65).

ized administration can exploit economies of scale in the designing of procedures, processing of returns, and employment of specialized auditors. However, the highly deconcentrated structure of the federal Ministry of Taxation currently does not allow exploiting all the potential economies of scale.

In addition, for taxes whose base is not shared with the federal government, there are not enough incentives for rigorous federal enforcement. The international experience shows some ways to provide those incentives for federal tax agents—for example, letting the federal administration retain a share of the collection proceeds.

Both the international practice and theory suggest that subnational tax administration can be more efficient in administering certain taxes that require specific local information. For example, information on the issuance of building permits by local authorities helps to schedule reassessment of property. Also, local government licensing of services provided by small businesses (such as retail operations) generates information that is helpful in the assessment of presumptive income. Thus, there are potential gains from independent subnational administration of taxes on property and presumptive income. Also, in theory, local authorities are better positioned for the establishment of taxpayer contact. However, the federal Ministry of Taxation also has local offices at every rayon/city center. Local tax autonomy would make a difference in this respect only at the settlement level. However, at that level, independent tax administration is likely to be unfeasible.

Russia has experience in property tax administration at the local level. In the original draft of the *Tax Code*, chapter 40 provided for the introduction of a Western-style real estate property tax at the regional level, which was to be shared with local governments. Two pilot cities (Novgorod and Tver) have been working on the introduction of the real estate tax since 1997.³⁵ The city of Novgorod created a municipal register of property specifically for taxation purposes. For every property object, the register contains data on its technical characteristics, legal status, and the holders of property rights. These data are used for computer-aided mass appraisal of the registered property objects. At present, for about 10 large enterprises that own the land on which their facilities sit, the two property taxes have been replaced by the new real estate tax. Although the property register covers all property objects, the present system of revenue sharing with the regional government reportedly discourages city officials from tapping this tax base. The idea is that at later stages the experiment will cover legal entities that lease land plots. The experiment in the city of Tver has fallen behind the one in Novgorod. Both experiments had to be completed by 2003 and the results were to be used in developing the federal legislation on this tax.³⁶ The results of the experiments along with the evidence presented in this paper would suggest that administration of the new real estate tax should not be centralized above the regional level.

If Russia chooses to allow independent subnational tax administration, certain safeguards would be necessary to prevent nonuniform application of the tax laws. For example, the federal Ministry of Taxation could provide comprehensive and uniform training to subnational tax officials. As well, the federal government should enforce strict adherence to audit guidelines. If possible, subnational tax administrations should use for each taxpayer the same taxpayer identification number as used by the central tax administration. In addition, the federal government should give subnational tax authorities access to federal databases to facilitate delinquency control.³⁷ The implementation of the real estate tax pilot project in Novgorod clearly exposed difficulties in accessing the databases of federal agencies. Discussions should be held between tax administration levels to provide for some type of information sharing, so that cross-checking activities can be performed. It also would be appropriate to form a group of subna-

35. Federal Law on Real Estate Tax Experiment in the Cities of Novgorod and Tver (110-FZ; July 20, 1997).

36. See footnote 21.

37. For instance, in Estonia revenue from the motor vehicle tax constitutes the majority of local tax revenues. In its present form, it has been imposed in Tallinn and in some smaller local government unit. However, matching the car's registered address and the residence of its owner is complicated. This is because cooperation with the Central Car Register Center is relatively hard to accomplish and thereby fines may be ineffective. (Högye and others 2000, p. 202).

tional tax administrators who could meet regularly to discuss common problems and solutions. This group, for instance, could prepare a model tax declaration so that each jurisdiction does not have to design its own form. It could also disseminate information on various issues, such as how to strengthen tax enforcement.

Appendix 5.3. Options Concerning Subnational Surcharges on Federal Taxes in Russia

As the Russian Federation continues its transition toward a new federalism, piggyback taxes (subnational surtaxes) have often been touted as a good solution to provide subnational levels of government with a certain level of tax autonomy. However, so far no single study has provided a comprehensive overview of what exactly piggyback taxes are, what conceptual issues should be considered before implementing a piggyback tax, which taxes are suited for piggyback collection and which are not, and what the international experience is in this regard. In this appendix, we define and consider the conceptual issues that should be addressed before implementing a piggyback tax and draw lessons and recommendations for the possible introduction of piggyback taxes in the Russian Federation.

What Is a Piggyback Tax?

Piggyback taxation—also known as subnational surtaxes—is an approach to taxation where regional or local governments “piggyback” (or establish a surtax) on an existing higher-level government tax. (See appendix box 5.3.1 for an example.)

A piggyback tax has three key characteristics:

- Subnational governments have discretion over the marginal tax rate to be charged within their jurisdictions.
- Subnational governments have no discretion over the tax base; the piggyback tax uses the same tax base as the national tax.
- The tax is administered and collected jointly with the higher-level tax by the higher-level government. The higher-level government collects the taxes for the lower-level governments and transfers the collections to the account of each lower-level government unit on a derivation basis.

These three characteristics combine to make piggyback taxes a unique approach to taxation that can be used in the context of a variety of taxes; for instance, one could piggyback sales tax, a personal income tax, and so on. Piggyback taxes could also be used at different levels in a fiscal federalist structure—a regional government could piggyback on a national tax; a local government could piggyback on a regional government tax; or local and regional governments could even both piggyback on a national-level tax.

Appendix Box 5.3.1 A Real-World Example

A real-world example of a piggyback tax may be helpful in illustrating the power of the piggyback approach as part of an overall tax system. Let us take the example of state-level sales taxes in the U.S. state of Georgia. Georgia has a state-level retail sales tax of 4 percent. However, in addition, state legislation allows local governments (which include county governments, municipal governments, and school-district governments) to institute a “local option sales tax” (LOST). If a local government chooses, it can incrementally increase the overall sales tax rate within its jurisdiction by one or more percentage points—up to a maximum of 7 percent. Once the Georgia Department of Revenue (DOR) is notified that a local government has adopted a LOST, the DOR will notify all affected taxpayers (that is, the retailers in the relevant jurisdiction) of the change in the overall tax rate. The DOR will then administer and collect all sales taxes (the state portion plus the local government portion) and remit the local share to each local government on a monthly basis. The DOR retains a small percentage of the collected local revenue as a fee for administering the piggyback sales tax.

Piggyback Taxation and the Principles of Revenue Assignment

The discussion on whether or not a country or a region should introduce a piggyback tax ought to take place in the context of the overall discussion of the design of the country's system of intergovernmental fiscal relations, and particularly in the context of the assignment of revenue sources among different levels of government within a country. Although an understanding of the basic architecture of tax systems and revenue assignments is crucial to the discussion of piggyback taxes, much has been written about the general principles of sound assignment of revenue sources. A complete review of the principles of revenue assignments would fall beyond the scope of this appendix.³⁸ Instead, this study will limit itself to the role that piggyback taxes could play within the context of a well-designed tax system and within the context of a properly conceived assignment of revenue sources. However, we should note that the implementation of piggyback taxes in Russia would first require resolving several important fiscal policy questions, including to what extent the federal government should involve itself in determining regional-local relations—for instance, to what extent the federal government can mandate and regulate the use of piggybacking between regional and local governments, or to what degree regional governments should be allowed to structure their own tax-sharing and coordinating mechanisms with local governments within their jurisdiction.³⁹

Consistent with the principles of revenue assignment, the overall attractiveness of piggyback taxes as a fiscal policy tool can be summarized by its main three benefits:

- Piggyback taxes allow a relatively high degree of centralized control over revenue sources.
- Piggyback taxes give subnational governments a certain degree of tax autonomy.
- Piggyback taxes are administratively simple and feasible.

ADMINISTRATIVE ARGUMENTS There are a number of very strong administrative arguments as to why a piggyback tax might be an attractive policy option in developed as well as developing or transitional economies.

First, piggyback taxation is an administratively simple and feasible option in the short run. The introduction of a piggyback tax allows for a certain degree of revenue autonomy to be achieved without having to introduce a separate subnational system of tax administration.⁴⁰ Administratively, tax authorities would have to make three modifications in the collections process to enable piggybacking:

- The tax administration would have to be able to determine in which subnational jurisdiction(s) each taxpayer resides for tax purposes. Of course, this requirement is the same for revenue sharing from a national tax.
- A separate line on the tax form would have to identify the surtax that is due to the regional or local government. However, for payment purposes, the full amount of the tax would be due to the national or higher-level authority.
- Administrative procedures would have to be established to clear revenue streams between jurisdictions for revenues that are collected in one jurisdiction but that should flow to the budget of a local government in another jurisdiction.

38. See McLure (1999) or Bird (1999) for a general discussion of the “revenue assignment question” in developing and transitional economies. Martinez-Vazquez and Boex (2001) provide a detailed discussion of revenue assignment issues in the Russian Federation.

39. Other key questions at the regional-local level include with which level of local government (rayons, settlements) revenues should be shared or piggybacked. This issue—clearly beyond the scope of this paper—is discussed in greater detail in appendix 3.1.

40. Of course, currently only the Federal Ministry of Taxation in Russia is empowered to enforce the collection of government revenues. Whether in the longer term there is a need for subnational tax administration falls beyond the scope of this appendix, but is addressed in another appendix to this chapter (appendix 5.2).

Second, a piggyback approach is an incentive-compatible collection mechanism. In the absence of a separate and independent subnational tax administration structure, one would have to devise a tax administration scheme under which the national tax administration has an incentive to collect revenues for subnational governments. After all, what incentives does the federal tax administration have to use its scarce resources to collect revenues for regional or local governments? A piggyback structure avoids this dilemma by tying the subnational tax liability inextricably to the taxpayer's national (higher-level government) liability. Thus the tax administration cannot help but collect the piggybacked amount if it wishes to collect its own revenues.

In order for a piggyback tax to function properly, it is important that the taxpayers' entire liability is due to the higher-level government, so that taxpayers have no option but to pay their entire (for instance, national plus subnational) tax liability. In fact, in the mid-1990s, for certain shared taxes (such as the enterprise assets tax) firms were able to pay their regional tax liability separately from their federal tax liability, and thus had the option to pay only one or the other depending on the incentives and the enforcement mechanisms in place. Thus, the key to incentive-compatible collection is to make sure that taxpayers only have a single tax liability, namely to the higher-level government, with the higher-level government responsible for passing on the proceeds from the piggybacked portion to the appropriate subnational authorities. Also, a further incentive might be given by allowing the national tax administration to retain a small percentage of the piggyback revenues as a collection fee, which would be earmarked for the tax collection office.

Third, a piggyback system is convenient for taxpayers. The compliance cost of a piggyback tax is extremely low compared to other own-source revenue mechanisms. The inclusion of a few extra lines on the higher-level tax form is typically sufficient to compute and collect the additional tax liability in the existing administrative process. Even when different taxes are well coordinated (for instance, a local income tax and a regional income tax), having an additional non-piggybacked tax results in administrative duplication and an increased compliance burden for taxpayers.

Fourth, while it is possible to achieve coordination between taxes at different levels of government that impose taxes on the same tax base, the desire to have a coordinated system often falls short of ensuring actual coordination between different levels of government. Failure to properly coordinate taxes at different levels of government may result in excessively high marginal tax rates, which could result in increased tax avoidance and tax evasion. Piggyback taxes would avoid this dilemma by allowing the higher-level government to specify a specific range of allowable marginal tax rates.

A fifth administrative advantage of piggyback taxes is that piggyback schemes are relatively transparent. For instance, while taxpayers would only have to make one payment, a piggyback property tax bill would clearly show how much of the collected taxes flows to each of the involved levels of government. Transparency is an important precondition to ensure accountability in the subnational budget process.

PIGGYBACK TAXES ALLOW FOR A HIGH DEGREE OF CENTRALIZED CONTROL OVER REVENUE SOURCES Although revenue sources were increasingly decentralized from the federal government to the regions in the early years of Russia's transition, since 1998 a recentralization has been taking place. There are inherent tensions between the desire of the central government to centralize control over revenue sources and the reduction of subnational tax autonomy. Failure to strike an adequate balance between the two would be a concern for the sustainability of the fiscal federalist structure.

On the one hand, based on the principles of revenue assignment, there are a number of legitimate reasons why many of the public sector revenue sources should increasingly be centralized. For instance, since national governments are typically responsible for ensuring macroeconomic stability, they often want to have tight control over key fiscal policy tools, including the

majority of tax instruments. Substantial revenue autonomy could jeopardize a national government's ability to use fiscal policy for macroeconomic stabilization purposes. Macroeconomic instability in the early years of Russia's transition certainly contributed to a view that federal control over revenue streams was increasingly important as a macroeconomic tool.

In addition, because the national government typically also carries out the redistributive function of government, it should be responsible for any progressive tax instruments (generally the personal income tax). For the same reason, major revenue sources that could cause horizontal fiscal imbalances should flow to the national as opposed to the subnational level. The large increases in regional fiscal inequalities that occurred in Russia in the 1990s should thus also have spurred an increased centralized assignment of revenue sources in order to reduce such horizontal fiscal imbalances.

In the context of the Russian Federation, it is also important to recall the recent history of the former Soviet Union, in that the collapse of the Soviet Union was brought about when the Soviet Republics refused to remit their tax collections to the Soviet Union. The political lesson—that control over the actual flow of resources continues to be politically relevant—is an important one and continues to be relevant in a climate with centrifugal forces pulling on the Russian Federation. Before a series of reforms in the “vertical power structure” were introduced by President Putin in 2000, the federal government lacked any effective means to control regions that failed to comply with federal regulations. While the tension between the federal and regional levels of government have subsided in favor of the federal government, ultimately the power of the federal government would be reduced if a significant portion of subnational revenues flowed directly to regional treasuries as own-source revenues as opposed to through the federal system. Excessive revenue autonomy would thus potentially allow recalcitrant regions to ignore federal tax legislation and regulations without the central government having any effective fiscal controls or incentives to force regional governments into compliance. Because piggyback revenue collections flow through the federal treasury system, piggyback taxes allow the federal government administrative control over resources while allowing regions discretion over tax rates.

PIGGYBACK TAXES ACCOMMODATE THE NEED FOR REVENUE AUTONOMY On the other hand, there are important implications for the performance of the system of intergovernmental fiscal relations that would result from a lack of revenue autonomy in a fiscal federalist system. Relating tax revenues to the benefits of public spending at the subnational level should be considered a key factor to ensure the optimal allocation of public resources in a decentralized system as well as a means to increase the accountability of subnational governments. In order to make sure that subnational government can provide the level of public services demanded by their constituents, subnational governments need to have the ability to affect at the margin the level of their revenues by choosing at least their tax rates. In addition, a closer correspondence between subnational taxes and subnational expenditures will ensure greater participation and oversight in the allocation of public resources at the subnational level. Thus, increasing taxing authority and reducing subnational governments' dependence on the revenue and transfer decisions of the upper-level government would increase both the efficiency and the accountability of subnational governments.

While simply providing subnational governments with “regular” revenue autonomy could (arguably, at least) reduce the ability of the national government to maintain macroeconomic control or reduce the ability of the national government to redistribute resources in an equitable manner, piggyback taxes do not require that the national government give up complete control over these taxes as a social or macroeconomic policy tool. A piggyback set-up allows the higher-level government to provide lower-level governments with revenue discretion within certain federally set rate limits. Thus, overall, an important benefit of piggyback taxes is that central government officials feel piggyback taxes are an acceptable mechanism to provide subnational governments with a certain level of revenue autonomy.

Alternatives to Piggyback Taxes

While we argue that piggyback taxes are an attractive taxation option, it is hard to assess the advantages and disadvantages of a piggyback tax unless one compares this approach to subnational taxation directly against other available policy options—of which there are essentially two:

- Using revenue sharing or intergovernmental fiscal transfers, where subnational governments have no control over the tax rate or tax base; and
- Providing subnational governments with “regular” own-source revenues, possibly with varying degrees of coordination between the higher-level and lower-level governments.

REVENUE SHARING AS AN ALTERNATIVE TO PIGGYBACK TAXES Shared revenues could be defined as higher-level government taxes or revenue sources that are collected by the higher-level government and then partially (or in some cases, fully) shared (often on a derivation basis) with lower-level governments. Revenue sharing can be based on fixed sharing rates (so that all lower-level government jurisdictions receive the same share of revenues collected within their jurisdiction) or on a differentiated sharing rate structure, in which different jurisdictions receive a different percentage of the shared revenue source.

While piggyback taxes have some of the same features as shared revenue sources, the former provide greater subnational revenue autonomy because, unlike the latter, piggyback taxes allow subnational governments a certain measure of discretion over the tax rate. Thus, if instead of introducing a piggyback tax Russia would increase its reliance on revenue sharing, this would further limit the level of revenue discretion of subnational governments in the Russian Federation.

Mechanisms for Coordinating Tax Administration at Different Levels of Government and the Issue of Subnational Tax Competition

On normative grounds, there are no reasons why several levels of government could not apply a tax to the same base. On the contrary, multiple use of the same base, if properly coordinated, has been found to simplify administration and reduce compliance cost (Musgrave 1998). Of course, the coordination mechanism would have to ensure that the combined marginal tax rate does not become so high that the tax provides incentives for tax avoidance or tax evasion.

There are four major mechanisms of coordination that accommodate more than one level of government using the same base (Break 1980, p. 34): cooperative administration, coordinated tax bases, tax supplements, and centralized administration.

Under *cooperative administration*, different levels of government legislate and administer their own taxes, but regular exchange of information occurs among tax administrators of different governments. Even when the tax base and rates do not coincide among governments, this kind of coordination can be highly beneficial as work done by one administration can generate revenue for another with little additional cost. For example, in the United States, the federal Internal Revenue Service (IRS) may inform a state about audit findings for an individual living in that state (Mikesell 1999, p. 512). (Similarly, that state’s Department of Revenues will alert the IRS with their audit findings.) However, the tax compliance brought about by the threat of audit by the IRS is considered so severe that it has been argued that, if the federal government of the United States were to eliminate its income tax, the states would have to do the same, as they would lack the information necessary to administer and audit their state income taxes adequately (McLure 1999).

Under a *coordinated tax bases* approach, while each level of government still legislates and collects its own revenues, the subnational government links its tax to some element in the structure of a tax defined by a higher-level government. For example, several U.S. states levy their state individual income tax on a taxpayer’s federal adjusted gross income, so that the state income tax form simply begins with an amount taken from the federal income tax form. Similarly, a number of

localities have introduced a local sales tax using definitions taken from their state sales tax (Mikesell 1999, p. 512). Even if other elements of the tax (exemptions, rates, and so on) differ, this kind of coordination reduces compliance costs, as only one set of records can be used for both taxes.

A more dramatic reduction in compliance and administration costs is achieved with *tax supplements*. Again, each level of government administers its own revenues (so the taxpayer still has to complete separate tax forms for each level of government), but the subnational government's tax is defined as a "supplement" to the higher-level government's tax by defining the lower-level tax liability as a percentage of the tax paid to the higher level.

The strongest coordination is achieved via *central administration* of subnational surtaxes; in other words, introducing a piggybacked tax. As already noted, under full piggybacking a lower-level government applies its own rate to the tax base used by a higher level of government, while all administration and enforcement is provided by the higher-level government. The taxpayer files a single tax form with the higher-level government, which records and remits collections to the lower-level government.

Tax competition raises certain concerns about piggyback and subnational taxes that target business activities, such as the corporate income tax (CIT), the enterprise assets tax (EAT), and sales tax. As the CIT and the EAT are paid to the locality where the company headquarters are located, local governments can try to attract companies by lowering the rates on these two taxes. The competition in regional sales tax rates can be also be aggravated by cross-border shopping. For example, the Leningrad region, which surrounds the City of St. Petersburg, has not introduced the sales tax, trying to attract retailers from the City. To be sure, some degree of tax competition and other forms of competition between subnational governments is healthy and ought to be encouraged. The question is whether some forms of tax competition are healthy and in the longer term contribute to a more efficient and vibrant subnational government. This is a complex issue and we will not try to address it here. It will suffice to say that the design of piggybacking systems can easily address the problem of harmful or unfair competition, if one exists, by introducing minimum tax rates that all jurisdictions have to use.

Policy Implications and Recommendations for Russia

The preceding material provides a comprehensive overview of the concepts and international practice with regard to the application of piggyback taxes. International experiences reveal a variety of possibilities in the use of piggyback taxes, but also highlight a number of limitations. Based on the conceptual underpinnings of tax assignment and international experience, what are the implications for the tax structure and assignment of revenue sources between different levels of government in the Russian Federation?

CLARIFYING THE LEGISLATIVE CONFUSION AND PROVIDING SYSTEMATIC MECHANISMS FOR REVENUE DISCRETION Transparency and clarity in the assignment of revenue sources are preconditions to ensuring accountability and efficiency in the subnational public sector. However, one of many sources of confusion is that many "federal" revenue sources are shared between all three levels of government and that within many regions of the Russian Federation, regional governments still continue to apply regulated sharing rates in sharing their portion of "federal" taxes with their local governments, thus further blurring the lines between actual revenue sources and intergovernmental fiscal transfers, and further complicating the issue of "ownership" over revenue sources in the system of intergovernmental fiscal relations. Local constituents simply cannot hold their local elected officials accountable if they do not know who ultimately has control over their public resources.

We believe that the introduction of selected piggyback taxes at various levels of the federal structure can contribute significantly to a more structured and transparent revenue system. In order to ensure compliance with the traditional principles of sound revenue assignment, we

believe that an important first step in reforming Russia's tax structure should be to restructure and clarify the legislation that determines what constitutes federal, regional, and local revenue sources. Although international practices vary in this regard, the ultimate goal should be to come up with an internally consistent approach to classifying revenues—something that appears to be lacking in the current system.

A reasonable approach might be to classify revenues according to which level of government sets the tax rate for the revenue source, or alternatively, to classify revenues according to which level of government receives the tax revenues. With regard to piggybacking, a consistent approach to classifying revenues would consider a piggyback tax as two separate taxes. For instance, under the current system in Russia, the federal share of the CIT should be considered a federal revenue source since the federal government sets the rate for this portion of the CIT. We believe that revenues from the regional share of the CIT should be accounted for as a separate tax and considered a regional revenue source, as regions have control over the tax rate and receive the revenues from this source.

PIGGYBACKING ON THE PERSONAL INCOME TAX Both the normative literature and international experience strongly support the possibility of modifying Russia's current federal PIT into a full-fledged piggybacked personal income tax, thereby instantly providing regional governments (and potentially local governments) with substantial revenue discretion.

In fact, in the late 1990s, a piggyback personal income tax for subnational governments gained acceptance in analytical and academic circles in Russia and was proposed as part of the 1999 budget.⁴¹ However, there is a regional aversion to a regional piggyback PIT—probably because the revenues from the federal personal income tax are currently assigned entirely to the subnational level (Martinez-Vazquez and Boex 2001).

Currently, revenues from the "federal" personal income tax are entirely allocated to regional governments, which can share the proceeds with constituent localities. However, subnational policy discretion and accountability would be significantly enhanced if regional (and possibly even local) governments could determine a regional or local personal income surtax in addition to the federally established tax rate within bounds determined by federal legislation.

Imposing a federally defined minimum on the piggyback income tax rate would arguably preclude a "race to the bottom" in which subnational governments would compete with each other to have the lowest PIT rates. While yielding fiscal discretion to subnational governments, the federal government should avoid a situation in which there are large variations in the effective subnational PIT rates, as such excessive variations (especially at the local level) would give taxpayers fiscal incentive to avoid (by suboptimally locating in suburban jurisdictions with lower tax rates) or evade taxes (by falsely declaring residency in a jurisdiction with a low subnational PIT rate).

On the grounds of its redistributive impact, the assignment of revenue from the PIT to subnational governments appears justifiable given that the Russian PIT is essentially a nonprogressive tax, charged at a flat rate of 13 percent of taxable income. However, conditions for the successful introduction of a piggyback tax, in addition to making sure that the subnational rate structure is nonprogressive, include levying the tax on a destination basis so that (consistent with the benefits principle) the revenue accrues to the jurisdiction in which the taxpayer resides. This would require the Russian tax administration to adopt a series of administrative procedures to clear piggyback PIT collections that are collected in one jurisdiction through the withholdings process but should accrue to another jurisdiction.

41. The assignment of 100 percent or the majority of revenues from a progressive PIT to subnational governments, as Russia did until recently, was highly questionable. Since the assignment of revenues was on a derivation basis, the redistribution of income from the progressive rate structure only affected the jurisdiction where the tax was collected. Richer jurisdictions got more revenues and poorer jurisdictions less than they would have under an equivalent flat-rate tax.

PIGGYBACKING ON THE CORPORATE INCOME TAX Offering conceptual arguments and administrative reasons, fiscal policy experts often recommend against levying a CIT at the subnational level, irrespective of whether the tax is collected as shared revenue source, an own-revenue source, or a piggyback tax.

The bottleneck that international experience has revealed in administering a subnational CIT concerns the establishment of a practical link between the amount of the CIT paid in each jurisdiction and the benefits from subnational government services enjoyed by the taxpayer. International attempts to rely on this benefit principle by using apportionment formulas have been shown to be cumbersome and often even altogether unfeasible at the local level. As a result, we believe that the local surcharge on the CIT that was introduced in Russia is not the best replacement for the abolished local tax on turnover (that is, the housing and utilities tax). A piggyback PIT would be a far superior choice.

PIGGYBACKING ON EXCISES AND OTHER CONSUMPTION TAXES Currently Russia has a regional retail sales tax (RST) and a federal VAT. The assignment of a destination-based sales tax at the subnational level also accords with the benefit principle. However, the current plans are to phase out the regional sales tax by 2004. The choice between the RST and the VAT and how these two taxes interplay in revenue assignments in Russia is discussed in the following appendix.

Given that the VAT in Russia is collected on an origin basis, the assignment of the VAT fully to the federal government is justifiable since the sharing or piggybacking of the VAT on an origin basis would result in substantial inequities and distortions between regions. Although there has been increased interest in the concept of a subnational VAT in federal countries around the world, we believe that a subnational VAT would inherently pose a variety of administrative complications that would require such a strong and well-coordinated tax administration that the option is not desirable in the Russian Federation at this point in time.

A piggyback subnational VAT is even less common mainly because of the apportionment problems that arise when the VAT is shared with subnational governments. Credits and liabilities for the tax arise in different jurisdictions and there is no simple way to attribute the tax on a derivation basis unless an explicit, and arbitrary, formula is used. In recent years, tax analysts have developed the principles for a subnational VAT that could piggyback (or not) on a national VAT. However, the implementation of this formula requires an advanced level of national and subnational tax administration. Until now, the only successful example of a subnational VAT has been in the province of Quebec in Canada. A regional piggyback VAT could be an option in the longer term in Russia after the tax administration system has been modernized and improved substantially.

In contrast, the assignment of the presumably destination-based regional sales tax to the subnational level in Russia accords with the benefit principle, particularly at the regional level. The absence of a federal sales tax (and the desirability thereof in the presence of a federal VAT) makes a federal-regional piggyback sales tax an illogical option. However, from the perspective of piggybacking, we have seen from the international experience that countries such as the United States are able to use piggyback sales taxes at the regional-local level. Conceivably, and if the RST at the oblast level had not been phased out in Russia, a regional-local piggyback sales tax akin to the LOST adopted in several U.S. states, while far from perfect, could have provided a possible mechanism to provide local governments in Russia with a certain degree of revenue discretion.⁴² Now that the RST has in fact been phased out, the issue is moot.

In addition to a possible piggyback sales tax at the local level, excises offer another potential source for federal-regional or even regional-local piggybacking. Excises are often suggested by

42. The coexistence of an RST with a VAT is less problematic as long as the RST rate is kept low. This would mean that, for example, the rayon and city governments could be given an optional rate of 1 percent piggybacking on the oblast sales tax, which at its height had a maximum rate of 5 percent.

fiscal policy experts as a source of significant revenue at the regional level because they are politically acceptable, easily administered at this level, and allow for rates differentiated by region. Moreover, the benefit principle accords with the assignment of (destination-based) excises on alcohol and tobacco to the regional level (to the extent that the latter is responsible for health-care) and on vehicles and fuel (to the extent that regions engage in road construction and maintenance).⁴³ However, it has been argued that a very elastic demand for healthcare should not be tied to the inelastic revenue from sumptuary excises (Bird 1999). There is a stronger economic case for earmarking vehicle-related taxes for expenditures on roads (Bahl and Linn 1992).

PIGGYBACKING OF PROPERTY TAXES The current subnational tax on enterprise property does not follow the benefit principle if the property value is weakly related to the capital services derived from these assets. In addition, in Russia, the enterprise assets tax is paid where the company is headquartered, not where the property is located. However, a reformed real estate property tax along the lines the federal government had supported with the pilot work in Tver and Novgorod would obey the benefit principle and provide adequate basis for subnational revenue autonomy.

The property tax is appropriate for assignment to the subnational level, as property value often reflects the benefits of public services (such as proximity to a metro station in Moscow). However, in transitional and developing countries, this tax often degenerates into a business tax, particularly for those businesses where it can plausibly be assumed that some or all of the tax is “exported”—that is, “someone other than local residents pays” it (Bird 1999). This clearly violates the benefit principle.

The current version of the *Tax Code* will assign the new real estate tax to the regional level. Much has been debated about the wisdom of this measure and why it may be better to assign the property tax, as is the case in most countries, to the local level. The merits of regional assignment versus local assignment will not be discussed here. Let us just say that the main argument in favor of a regional (oblast) assignment is that the complexity of developing and updating a fiscal cadastre with market values of real estate property may be well beyond the capabilities of many cities and rayons. Without judging the veracity of this statement, we can make a clear case in favor of local surcharges on the regional real estate property tax. These local surcharges would use the same base and property assessment as the regional tax and would impose a flat rate surcharge. As in the discussion on the PIT, the property tax piggyback rate for local governments could have minimum and maximum rates.

Appendix 5.4. The Value-Added Tax and the Retail Sales Tax in the Russian Federation

In 1998, a sales tax was reintroduced in the Russian Federation.⁴⁴ Previously, regional governments had the authority to levy a sales tax of up to 5 percent.⁴⁵ The introduction of a sales tax was a government compromise aimed at appeasing regional governments when the federal government reduced their share of the VAT. At some point in the 1990s, the federal government also planned to further reduce the VAT rate while increasing the sales tax rate for goods subject to excises up to 10 percent. This proposal was ultimately rejected within the government. However, in January 2004 the VAT rate was reduced from 20 to 18 percent and the regional sales tax was eliminated altogether at the same time. This follows the 2003 elimination of the road user fee, the last of the old

43. There used to be a tax-exporting problem with excises on alcohol in the Russian Federation when these were levied at production. Now that only 50 percent of this excise is levied on production and the other 50 percent on the wholesale of alcohol, the extent of tax exporting is believed to be less.

44. Parts of this appendix build on United States Government Technical Assistance Team (1998).

45. This right was technically rescinded in 1996, but some regions continued to impose a sales tax until at least 1997.

generation of turnover taxes (of which there were many), which used to provide the bulk of own-source revenues for subnational governments. As a result of these measures, the extent of revenue autonomy of local governments dropped to an unprecedented low level. This situation may force the central government to revisit the issue of subnational sales tax in the future.

Thus the Russian Government will have to face an important tax policy decision. This decision involves the structure and composition of indirect taxation, revenue assignments between the federal government and subnational governments, and most important, the choice of the general consumption tax between the VAT and the RST. The current VAT is levied at a general rate of 18 percent, and it is assigned fully to the federal government's budget. The tax base of the VAT is relatively broad, but there are some exemptions that, as explained below, could cause difficulties in terms of administration. Before being eliminated, the regional sales tax had been rationalized in the revisions of chapter 27 of the *Tax Code*, effective January 2002. The previous law that governed the regional sales was somewhat ambiguous regarding the definition of the tax base, and many regional governments had seized the opportunity to define tax bases that suited local needs. Constitutional challenges brought by business firms in 2001 resulted in a tax law that was more specific in terms of tax base, methods of transaction, and the implications of nexus.⁴⁶

At issue are whether the RST at the regional level has a legitimate role when a general consumption tax, the VAT, already exists at the federal level; whether the VAT or the RST is a "better" tax for regional governments; and whether Russia's decision to eliminate the sales tax at the subnational level is justified. Russia's case of a national VAT and subnational RST is relatively unique. The current system in Canada is the most closely aligned system of dual (but different) consumption taxes, including a VAT at the federal level and a sales tax (as well as a provincial VAT in Quebec) at the subnational level. Other countries, such as Nicaragua, have contributed to the international exception of a cohabitation of a national VAT and a subnational sales tax. In the case of Nicaragua, the subnational sales tax was eliminated recently at the insistence of the IMF.⁴⁷

Both taxes have significant merits, and in fact, if they are imposed in their purest forms, they are equivalent. The actual definition of the tax base and the administration of the taxes tend to create differences between these taxes. Based on these practical components of each tax, one might be preferred over the other. We do not believe that there is a clear-cut winner at the subnational level.

The purpose of this appendix is threefold. First, we provide a background analysis of the merits of the VAT in relation to the RST from the perspectives of best principles and international experience. (An analysis of the relative merits and problems of each tax must include the economic impacts, distributional impacts, and administration of the tax, as well as related intergovernmental issues.) Second, we examine in the current Russian context the pros and cons of relying on the VAT or the RST as the main consumption tax from the perspectives of tax policy, tax administration, and intergovernmental fiscal relations. Third, we provide several policy recommendations on this issue.

A Comparison of the Value-Added Tax, Retail Sales Tax, and Turnover Tax

The value-added tax (VAT) and the retail sales tax (RST) are both taxes on consumption. The turnover tax is a simpler form of taxation in which all transactions are taxed. The turnover tax quickly results in "cascading" with superimposed increases in prices at each stage of production and consumption. The VAT is far and away the most popular of the two taxes, with 123

46. The new law prevents a region from imposing a sales tax on purchases made in a non-sales-tax region. This is always a complex issue. For example in the United States, the "use principle" is an attempt to impose tax in the jurisdiction in which the good is put into use if the good is purchased out of state.

47. The "orthodoxy" is that there is no room for the RTS and the VAT to exist side by side because given the different structure of the two taxes both compliance and administration costs increase quite considerably.

countries imposing it in 2001 (Ebrill and others 2001). By contrast, the RST is used widely in only a few countries including the United States, India, and Canada.

As noted above, the VAT and RST are theoretically equivalent taxes. In practice, the taxes have different impacts on production decisions and prices—and therefore on revenues. If a VAT works well (in terms of its invoice-credit mechanism), it is less easily evaded than the RST. The impact of this “self-policing” of the VAT is diminished when the invoice-credit system does not work well (that is, credits are not paid on time if paid at all) or if the invoice-credit is not extended to the retail level. The RST, properly administered, also leaves a trail of suspension certificates (see below), but from a tax enforcement viewpoint the trail left by the invoice-credit system under the VAT is more effective.

The VAT is also easier to remove on producer goods, since this mechanism is built into the system. In the case of the RST, producers would have to be specifically exempted from the system in order to avoid the potential for tax cascading. Under the RST, it is necessary for the supplier to determine whether a good is being used for private consumption or as an input in a production process. This discretion can result in the taxation of inputs under the RST or it can open the way for negotiation and potential corruption of the tax administration.

The differences between the VAT and the RST are largely the result of their administration. Both systems have an element of self-policing. Under the RST, the vendor polices the system by accepting a suspension certificate, which is an approval of the tax administration to exempt the purchaser from sales tax. Under a VAT system, the tax authorities police the system. But to take advantage of the anti-evasion benefits of the VAT, the tax administration must be able to process VAT returns and ensure that businesses do not receive credits for tax not paid, and for tax paid on purchases for household use. Businesses would have to remit taxes to the authorities for both systems, but the number of such remittances would be greater in the case of the VAT owing to its nature of collection at each stage of the production process.

In the recent past, a relatively high degree of barter and nonmonetary transactions in Russia’s economy arguably affected the administration and evasion of the VAT versus the RST. One of the original cases made for a retail sales tax in Russia is that retail (final consumption) transactions are less likely to be in the form of barter or other nonmonetary methods of payment than producer or wholesale transactions. This suggests that the RST may be better able to bring in cash revenues versus in-kind revenues to the budgets of regional governments. Assuming this is the case, does an RST have an advantage over the VAT? Under the extreme scenario of a system in which all transactions except final sales are barter and no VAT is paid on any intermediate transaction, both the VAT and the RST would collect the same (cash) revenues. Clearly, given equivalent tax structures, the RST cannot collect more revenues than the VAT. However, if nonmonetary forms of tax payments are allowed (and again assuming that these are more frequent at intermediate stages of the production), the RST can raise a higher proportion of tax revenues in cash (as opposed to nonmonetary forms).⁴⁸

The Nature of the Value-Added Tax, the Retail Sales Tax, and Turnover Tax

The value-added tax (VAT) and the retail sales tax (RST) are, at least in theory, alternative ways of taxing sales to consumers. By comparison, a turnover tax applies to all (nonexempt) sales, including those to business. A tax on consumption must be applied only to sales to households; it cannot be applied to sales to business, including sales of capital goods. The reason for the

48. Note that in the case of barter transactions where the VAT is paid, it could also be fully credited. Therefore no distortions should be introduced into the prices of intermediate goods. Under current Russian Federation legislation, barter deals are regarded as supplies for VAT purposes with valuation of the transaction based on average selling prices/open market value, and, in theory, credit may be taken on barter transactions.

exemption, as well as how the VAT and RST use different mechanisms to achieve this objective, is explained below.

- *Retail sales tax.* An ideal RST achieves the objective of exempting sales to business directly; it is applied only to sales to households and exempts all sales to businesses. In practice all extant RSTs are imposed on many sales to business, which increases the price of inputs in the production process. Note that this is a matter of administration and structure, not of the economics of the tax.
- *Value-added tax.* The VAT achieves taxation of consumption in a different manner. It is collected on virtually all sales. Taxation of business purchases is eliminated by allowing registered traders to deduct (take “input credit” for) taxes paid on purchases from tax due on sales. Thus the only tax that is not eliminated by input credits is that paid by consumers.
- *Turnover tax.* As noted, the turnover tax, by its very nature is imposed on many sales to business and thus violates the basic objective of taxing only sales to households. It is, by design, a cascading tax that can generate large revenues.

The Mechanics of Taxation under the Three Systems

To clarify the mechanics of the three systems, assume that production involves three stages, A, B, and C (which might be manufacturing, wholesaling, and retailing), as shown in appendix table 5.4.1.⁴⁹ The top panel provides a benchmark; it is used initially to illustrate the results of

Appendix Table 5.4.1 *Comparison of the VAT, the RST, and Turnover Taxes*

Panel I. Benchmark case

<i>Stage of production</i>	<i>Purchases</i>	<i>Sales</i>	<i>VAT on sales</i>	<i>VAT on purchases</i>	<i>Net VAT</i>	<i>RST</i>	<i>Turnover tax</i>
1a. Manufacture	—	300	30	0	30	—	15
2a. Wholesale	300	700	70	30	40	—	35
3a. Retail	700	1,000	—	—	$(1,000 - 700) \cdot 0.1 = 30$	100	50
4a. Total Tax liability	—	—	—	—	100	100	100

Panel II. Vertically integrated production

<i>Stage of production</i>	<i>Purchases</i>	<i>Sales</i>	<i>VAT on sales</i>	<i>VAT on purchases</i>	<i>Net VAT</i>	<i>RST</i>	<i>Turnover tax</i>
1b. Manufacture	—	—	—	—	—	—	—
2b. Wholesale	—	—	—	—	—	—	—
3b. Retail	—	1,000	—	—	$(1,000 - 0) \cdot 0.1 = 100$	100	50
4b. Total tax liability	—	—	—	—	100	100	50

Panel III. Value added in early stages

<i>Stage of production</i>	<i>Purchases</i>	<i>Sales</i>	<i>VAT on sales</i>	<i>VAT on purchases</i>	<i>Net VAT</i>	<i>RST</i>	<i>Turnover tax</i>
1c. Manufacture	—	800	80	0	80	—	40
2c. Wholesale	800	900	90	80	10	—	45
3c. Retail	900	1,000	—	—	$(1,000 - 900) \cdot 0.1 = 10$	100	50
4c. Total tax liability	—	—	—	—	100	100	135

Source: Authors.

49. The table assumes that there is no input tax shown at the manufacturing stage although there is likely to be some deduction on raw materials.

the three taxes and then to compare the results under other assumptions about the production-distribution process. Appendix table 5.4.1 also shows a situation of complete vertical integration, in which all production and distribution occur in one stage (second panel), and another situation in which value is added earlier in the production-distribution process (third panel).⁵⁰

In all three cases total sales to consumers are the same: RUR1,000.

- *RST*. Under a 10 percent RST, RUR100 would be obtained on RUR1,000 of sales to consumers at the retail stage. This is true regardless of the nature of the production-distribution process, as long as all sales to consumers, and only those sales, are subject to tax. To verify this, compare the RST columns in each case of appendix table 5.4.1.
- *VAT*. Under a 10 percent VAT, each firm would pay a tax of 10 percent on its sales, but would receive credit for tax paid on its purchases; see appendix table 5.4.1's columns labeled VAT on purchases, VAT on sales, and Net VAT. In the aggregate, the tax paid under the VAT is the same as under the RST. Again, this result is independent of the production-distribution process. Because credit is allowed for VAT paid on purchases, as shown on invoices, this method of implementing a VAT is commonly called the credit-invoice VAT.
- *Turnover tax*. Under the conditions assumed in the top panel (benchmark case) of appendix table 5.4.1, a 5 percent turnover tax would produce the same revenue as the RST and the VAT; this is shown in the Turnover tax column of the table. If, however, the production-distribution process were different, aggregate tax liability could be different. In the case of complete vertical integration, a 5 percent turnover tax would produce revenues of only RUR50, instead of RUR100; see the second panel of appendix table 5.4.1. By comparison, if value were added much earlier in the production-distribution process, as in the bottom panel of the table, a 5 percent turnover tax would yield revenues of RUR135—more than the revenue raised from a 10 percent tax on sales to consumers.

More on the Administration of the Value-Added Tax and the Retail Sales Tax

Having outlined the basic framework of these taxes above, we now turn to a discussion of specific administration issues.

THE PROBLEM OF ELIMINATING TAX ON BUSINESS PURCHASES The RST suffers from a basic administrative problem: how to achieve the objective of exempting sales to business (by suspending taxation of business purchases), without opening the door to evasion by households claiming to make business purchases.⁵¹ (In addition, there is the political problem of public pressure not to exempt all sales to business, thus converting the RST into a turnover tax.) Attempting to exempt all purchases by business places the vendor in the unenviable position of determining whether each sale is taxable or legally exempt and creates an incentive for the vendor to “look the other way” when a household purchase masquerades as a business purchase. This problem can be ameliorated by the issuing of “suspension certificates” by the tax authorities. Such a certificate identifies the purchaser as a registered business eligible to make exempt purchases, and the business use of the item that makes it exempt. In this case vendors do not have to make the decisions themselves regarding what purchases ought to be exempt. However, auditors still must trace exempt sales to the purchaser to determine whether they are for legitimate business expenses.

50. For simplicity's sake, appendix table 5.4.1 assumes that the sale price will remain at RUR1,000 in the case of exemption at the pre-retail stage. It is likely that the sale price (and also the VAT due) will actually increase if the wholesaler wishes to maintain its profit level since its costs of purchase will be higher by the value of the exempted VAT.

51. Using the terminology of the draft Russian law authorizing oblast sales taxes, the question is how to distinguish legally taxable “sales to the public” from other (legally untaxed) sales.

By comparison, under the VAT, since all sales are taxable, the vendor does not need to make the distinction between taxable and exempt sales, but must account for tax paid on purchases, in order to claim input credits. Although it operates indirectly to remove tax on business inputs, for both administrative and political reasons, the VAT achieves the objective of avoiding taxation of sales to business much more thoroughly than does the RST.

The turnover tax does not have the benefit of removing taxes on business transactions through invoice-crediting (as with the VAT) or the use of suspension certificates (as with the RST). Taxing each transaction results in cascading, with increases in prices at each stage of production that multiply the impact of the tax applied in previous stages.

EXEMPTIONS AND ZERO-RATING The concept of exemption is straightforward under either the RST or the turnover tax—the sale of a particular good or service is simply not taxed. The situation is not so clear under the credit-invoice system. One must distinguish between exemption and zero-rating based on the stage of the production-distribution process (retail or pre-retail). If a sale is *exempt*, there is no tax, and no credit is allowed for tax paid at prior stages of the production-distribution process. By comparison, if a sale is *zero-rated*, there is no tax, but credit is allowed for tax paid at prior stages.

If a retail sale is exempt, it nonetheless bears a VAT from prior stages. Moreover, if a pre-retail sale is exempt, the aggregate tax on the retail sale is greater than in the absence of exemption, not less. By comparison, zero-rating of retail sales eliminates all tax, and zero-rating of pre-retail sales has no effect on the aggregate amount of tax on retail sales; in either case the aggregate tax is determined by the tax rate imposed at the retail stage. These results are illustrated in appendix table 5.4.2.

ADMINISTRATION AND COMPLIANCE COSTS OF DUAL CENTRAL VAT/SUBNATIONAL RST SYSTEMS Both the RST and the VAT involve considerable administrative and compliance costs, especially for small businesses. Clearly, imposing both systems on business increases their com-

Appendix Table 5.4.2 *Illustration of Exemption and Zero-Rating under a Credit-Invoice VAT, for Retail and Pre-Retail Sales*

	<i>Stage of production</i>			<i>Total</i>
	<i>A</i>	<i>B</i>	<i>C</i>	
<i>Exemption of retail stage</i>				
Tax on sales	30	70	0	100
Input credit	—	30	0	30
Net tax liability	30	40	0	70
<i>Zero-rating of retail stage</i>				
Tax on sales	30	70	0	100
Input credit	—	30	70	100
Net tax liability	30	40	-70	0
<i>Exemption of pre-retail stage</i>				
Tax on sales	30	0	100	130
Input credit	—	0	0	0
Net tax liability	30	0	100	130
<i>Zero-rating of pre-retail stage</i>				
Tax on sales	30	0	100	130
Input credit	—	30	0	30
Net tax liability	30	-30	100	100

Source: Authors.

pliance costs significantly. Because the two systems operate quite differently, implementing both also increases tax administration costs significantly.

How much compliance and administration costs are increased by a dual VAT/RST system depends crucially on how different the bases of the two taxes are. These costs are substantially higher if the exemptions provided by the RST and the VAT are not substantially identical. In that case, vendors would need to distinguish between four possible combinations of tax treatments: taxable under both taxes; exempt under both; and taxable under one, but not the other. To minimize this type of complexity, the bases of the two taxes should be coordinated so that a given sale to households is either taxed or exempted under both taxes.

Economic Effects of the Value-Added Tax, the Retail Sales Tax, and Turnover Tax

VAT. Because the VAT is, in effect, levied only on sales to consumers, it is essentially neutral with regard to most economic decisions.⁵² That is, it does not affect the cost of capital and thus does not distort the choice between saving and consuming, the capital-labor ratio, or other production decisions. It does, of course, raise the price of consumer goods. This is an important point to understand because of the widespread perception in Russia that the VAT is paid by businesses and that this type of tax distorts business decisions.

RST. A pure RST (one that is applied only to consumer goods) would have the same effects as a VAT. In fact, in the real world RSTs are far less pure than VATs. It has been estimated that about 40 percent of RSTs levied by the states of the United States and some of the provinces of Canada fall on sales to business, including many capital goods. Thus they distort many choices: the saving/consumption choice, the capital-labor ratio, and production decisions. (See also the discussion of the effects of turnover taxes, which an impure RST resembles). Moreover, unlike the typical VAT, RSTs typically exempt many services provided to consumers, thereby distorting consumer choices.

Turnover taxes. Turnover taxes are notorious for their distortions of economic decisions. The degree of distortion of saving-consumption choices and capital-labor choices depends on whether or not the tax applies to capital goods; these distortions are greater, the more capital goods are taxed. What is clear is that turnover taxes encourage vertical integration in order to avoid paying tax several times as a product moves through the production-distribution process. The resulting concentration of industry is one thing Russia can ill afford.

Intergovernmental Fiscal Relations and the Choice between the VAT and the RST

Two general principles to keep in mind for the proper design of a system of intergovernmental fiscal relations have a significant bearing on the choice between the VAT and the RST. These principles are the revenue sufficiency (or vertical balance) between federal and subnational governments, and the revenue autonomy of subnational governments.

The VAT is just about the poorest choice to fulfill either of those two principles, and this seems to have been heeded by the Russian Federation. The introduction of regional VATs to provide regional tax autonomy can easily lead to chaotic situations, as the experience of Brazil has shown in recent years. In pure technical terms, the VAT may also be a poor tax to be shared with subnational governments on a derivation basis. The crediting and debiting of VAT in different regions make it absolutely unsuitable for direct revenue sharing on a derivation basis. Russia's federal government used to share VAT revenues with the regions. Besides the arbitrariness of crediting and debiting the VAT across regions, the tax was not truly shared on a derivation basis (that is, according to where the tax is generated); rather it was paid according to the place of registration of business firms.

52. This is the case for a consumption type VAT, where all VAT paid on intermediate inputs, including capital goods, is creditable.

In the past few years, the potential for a subnational VAT has attracted more positive attention than in previous decades. New legislation in India has mandated the development of a subnational VAT, although few specific plans have yet been made to institute a subnational VAT. The advent of the European Union has led to unique opportunities to develop a VAT that crosses international borders, although it should be noted that the European Union is a very different context than that of the federation of Russia.⁵³ Canada's system of harmonized sales tax and Quebec's sales tax are two examples of the coexistence of sales and value-added taxes. In these cases the federal government plays an important role in assisting or carrying out tax administration. In the case of Quebec, there is no attempt to tax interprovincial purchases made by end-consumers (Bird and Gendron 2001). While Canada, Brazil, and Argentina offer some examples of the potential to impose regional VATs, in most respects a VAT at the subnational level would entail extensive administrative requirements and high administrative costs.

Because of these problems, there is a very strong case to be made in favor of the full assignment of VAT revenues to the federal government, as is currently done. This was not an easy political decision for Russia, but now that the VAT is fully assigned to the federal government, the issue of revenue sufficiency becomes even more important for subnational governments.⁵⁴ If full assignment were changed, the second-best alternative would be the sharing of the VAT with the regions based on simple formulas.⁵⁵

A subnational RST, on the other hand, is in theory an adequate instrument to fulfill the two principles of revenue sufficiency and revenue autonomy for subnational governments. An RST can be a reliable source of revenue for subnational governments and it does not require a high degree of coordination among those governments. In theory, an RST could also be more easily shared among different levels of government. The problem with this solution lies in the costs of administering and complying with a regional RST when a VAT already exists at the federal level. In addition, there are many other reasonable alternatives besides an RST for providing revenue autonomy to subnational governments. One such alternative used in many countries and often recommended in Russia is a subnational government personal income tax with some degree of rate discretion that piggybacks on the federal personal income tax. This option appears less attractive than the RST option because in Russia subnational governments have historically been assigned almost the entirety of revenues collected from the personal income tax, although there have been changes to this intergovernmental arrangement in the past several years.

The introduction of an intergovernmental fiscal-relations perspective in the choice between the VAT and the RST reveals the existence of worthwhile but also conflicting objectives and the possible need for a compromise. While the VAT may be superior to the RST from economic-efficiency and tax-administration viewpoints, the goals of revenue sufficiency and revenue autonomy for subnational governments cannot be fulfilled with the VAT choice under the current tax administration in the Russian Federation.

Advantages and Disadvantages of the VAT and the RST in a Russian Context

The comparison of the VAT and the RST in the previous section raises a number of issues that are generally applicable to the Russian context. However, the peculiarities of Russian tax legislation and of the country's transition economy require caution in the direct application of conclusions from other countries. Hence it is useful to summarize the main reasons for and against the use of the VAT in relation to the RST in the Russian Federation. (We omit any further dis-

53. Keen and Smith (2000) specifically analyze this issue as it pertains to the European Union.

54. Early attempts at full federal assignment with the VAT were not successful. This assignment was tried during 1991–92, in the *Law on the Basic Principles of Taxation* yet regions failed to comply with the law and the political system accommodated itself to this refusal.

55. For example, the federal government in Germany shares VAT revenues with the regions on a per capita basis.

cussion of the turnover taxes here since they have been demonstrated to be an inferior choice; as a matter of fact, the only existing policy proposals are for their elimination.)

Advantages associated with the use of the VAT in Russia are as follows:

- The VAT works. It brings in approximately 45 percent of federal tax revenues. The VAT has become the workhorse of Russia's tax system.
- Taxpayers and tax inspectors understand the tax principally because it has been in operation for over a decade. This is relevant given the poor record in taxpayer education and tax inspector training.
- Use of a VAT is a precondition for entry into the European Union. While this may not be seen as an issue for Russia at present it is clear that many of the former Central and Eastern European countries see EU entry as the next stage in their economic reforms. Given that the Russian Federation already has a VAT that is moving toward the EU style of operation by adopting many of the EU principles for a VAT, it is not appropriate to seek to move away from it as the primary revenue-raising form of taxation.
- The tax invoice system of the VAT offers an excellent audit trail. From a self-policing perspective, there is a real incentive for the buyer of goods and services to demand an invoice in order to secure deduction of input tax, even though this link currently is weak in Russia because the invoice-credit system does not extend to the retail sector.
- Import VAT is relatively easy to collect. Even though there may be evasion further along the supply chain, a VAT imposed at import ensures that the first point of entry into the country sees tax collection. Reducing the VAT rate in exchange for a higher RST rate will sacrifice this stable collection point and reduce revenue yields.
- In theory, revenue is turned over to the government on a regular basis owing to the taxation at each stage of production.
- The value-added concept of the tax ensures that even though there may be suppression of sales by the retailer, much tax collection already will have taken place at earlier points in the chain. This is important in Russia because of the existing high levels of evasion.
- It can also be argued that the VAT improves accounting standards. This is particularly seen in the instructions introduced by the Government (Decision No 914 of 29 July 1996) that require VAT payers to maintain detailed purchase and sales books supplemented by tax invoices in order to support VAT declarations.
- The VAT covers goods *and* services. International experience suggests that an RST is not likely to offer the same level of coverage simply because of the conventions of the tax, not because of the economics of the tax.

Disadvantages associated with the current VAT in Russia are as follows:

- Clearly no further improvements can be made to bring the Russian VAT system up to international standards. There are still some distortions caused by limitations on input tax credit owing to exemptions and some restrictions on capital purchases.
- The sharing of the VAT with regions on a destination basis is inappropriate because of the uneven crediting and debiting of the tax across regions. It is difficult to envision a regional-level VAT under the current state of tax administration in Russia.
- Apportionment of VAT revenues is actually based on the place of registration of businesses. The current system penalizes regions with significant export activity and is non-equalizing.
- The VAT may not be an economically neutral tax in Russia. For example, the crediting of the VAT on previous inputs might not have worked well owing to significant barter exchanges occurring between firms in the recent past, and the high likelihood that no liability is incurred in barter transfers between firms. This can result in distortions and higher prices for outputs.

- The tax administration does not currently monitor transactions and invoices, thereby increasing tax evasion and lowering the revenue collected.
- The tax administration is such that enforcement advantages of the VAT (invoice-crediting) are not achieved. In particular, the State Tax Services do not appear to have taken advantage of the verification opportunities offered by the invoice-based VAT.
- The integrity of the invoice-credit system is at risk because of late payment of credits. This delay in credit payments is also an increased cost of production that is not foreseen in a VAT in which credits are immediately returned in the form of refunds.

Advantages of the RST are as follows:

- At moderate rates, the RTS can provide valuable revenue autonomy for subnational governments.
- If the RST has the same structure of exemptions as the VAT, the tax administration and taxpayer compliance costs associated with the RST can be greatly reduced.
- If retail transactions are more likely to be carried out in cash than are intermediate transactions, the RTS can raise a higher proportion of revenues in cash rather than in barter and other nonmonetary means of payment that are used with the VAT.
- If businesses are properly exempted from the tax, it does not impose an undue burden on the production process.
- At relatively low rates (5 percent), there is less reason for the vendor to defraud the government.
- With its relatively low rates, the tax is suitable for the subnational level. The efficiency is enhanced when similar tax bases are used among governments.

Disadvantages associated with the RST are as follows:

- In practice, the RST is levied on business purchases as well as on final consumption, which affects economic efficiency.
- The retail level is noted as the “weakest link” in tax administration owing to the vast number of retail outlets and the ability for retail sales to be made without official documentation (“under the table”).
- If the RST were to become a major revenue producer in Russia, there would have to be significant investment in tax administration to enforce this new tax properly. To date, the concentration has been on VAT administration and not on RST administration.
- The RST does not offer the same level of audit security as the VAT, especially after the extension of the invoice-credit system under the VAT to the retail sector.

Policy Recommendations

There are a number of policy options that address important tax-related issues for the Russian Federation such as revenue yield, administration, and intergovernmental issues. The discussion above demonstrates that both the VAT and the RST have strengths and weaknesses, particularly in the context of Russia. A high RST rate is not feasible. However, as instituted at the 5 percent rate, the RST can be a useful tax for regional governments. The policy option that was considered during 1998–99 of allowing an RST rate of 10 percent would put Russia’s RST outside the envelope of international norms for such a tax and has been estimated to yield a third less than a VAT of the same rate. A 10 percent RST rate would likely lead to significant tax evasion. In addition, a high rate would exacerbate many of the RST’s identified shortcomings.

The policy option envisioned in the Government’s draft *Tax Code* and the stabilization package of the summer of 1998, proposing a sales tax for regional governments of up to 5 percent levied on retail (final) transactions coexisting with the federal VAT, was a reasonable and acceptable compromise between the conflicting goals confronting the Russian Government.

Even though the existence of a regional sales tax side by side with the federal VAT could somewhat increase administrative complexity and compliance costs for taxpayers, these additional costs are acceptable in light of the goal of providing regional governments with revenue autonomy. The use of both a sales and a value-added tax has precedence in some countries (notably Canada). The additional costs in tax administration and taxpayer compliance can be significantly reduced if the exemptions of both the federal VAT and the regional RST are made exactly the same.

The 1998 sales tax law was not without its problems and some of them had been alleviated in the new sales tax legislation that is chapter 27 of the *Tax Code*. However, there were still some problematic issues, including the following:

- There were a number of exemptions—which opened the way for tax evasion, tax base erosion, and equity problems in the system. Among these were certain types of services, children’s clothing, and periodicals.
- The *Code* did not require that the retail sales tax be added to the price of taxable purchases at the point of purchase. Requiring this and also requiring that invoices or receipts provide the amount of the purchase before the application of tax, the amount of the tax, and the amount of the purchase with tax included would have increased the transparency of the tax.
- The *Tax Code* should have placed the responsibility for collection on the vendor and not on the purchaser. The vendor was part of the tax compliance system and making the vendor responsible would have eased the administration of the system.
- The *Tax Code* should have defined retailer liability as the value of all taxable product sold during the reporting period (including the value-added tax and excises for goods sold subject to the excise tax), not including the tax itself, multiplied by the retail sales tax rate applicable for those sales. This was not made clear in the current *Code*.

This policy formula does not offer a solution to the problems of assigning VAT revenues fully to the federal government or—the other side of the coin—providing subnational governments with sufficient substitute revenues (for the sharing of the VAT) and more revenue autonomy. The Government could explore a further reduction of the VAT rate at the federal level. The rate of the regional RTS could be set at 6 percent but any higher rates would likely promote tax evasion and complicate the RST’s administration with the federal VAT. Another option is to return to sharing the federal VAT with the regions but not on a derivation basis. A simple apportionment formula, such as per capita basis or an estimate of regional private expenditures, should be used instead.

As noted earlier, the VAT is generally a more efficient (economically neutral) tax than the RST because of tax practices, not because of the basic structure of the two taxes. Both are far more neutral than a turnover tax. Reducing reliance on the VAT and increasing reliance on the RST would, on balance, increase the distortionary impact of the Russian tax system; by comparison, replacing turnover taxes with the VAT would reduce distortions. If the RST replaced revenues from the VAT, there would be somewhat more pressure for vertical integration in order to avoid tax on sales to business, unless businesses were successfully exempted from the RST. Such pressure would be reduced if the VAT replaced the turnover tax. Revenues could, of course, be comparable under any of the taxes; that is a matter of adjusting tax rates to be revenue neutral. However, while a relatively low RST rate could be competently administered, rates higher than 5–6 percent are likely to result in significant tax evasion.

As noted above, imposing both an RST and a VAT does impose additional costs because it subjects taxpayers to the requirement of complying with two different systems and requires a duplication of administrative effort. However, a dual system can exist, in particular where the VAT is a federal government tax, the RST is a subnational (regional) government tax levied at a relatively low rate, and exemptions under the VAT and the RST are essentially the same.

6

Intergovernmental Fiscal Transfers

As in many other countries, expenditure assignments in the Russian Federation are not well matched with revenue assignments. Broadly speaking, a vertical imbalance exists when the revenue sources assigned to *each level of government* do not correspond to their expenditure responsibilities. Thus, as noted in chapter 5, tax revenues assigned to Russia's local governments, both from own-sources and shared revenue sources, cover slightly more than half of their expenditures. By contrast, horizontal imbalances arise from existing fiscal disparities *across subnational jurisdictions*. In virtually every country, the system of intergovernmental fiscal transfers plays an important role in balancing the relative expenditure needs of the different government levels with the resources available to each level of government (and also to different jurisdictions at the same level of government).

Recognizing the importance of the role of intergovernmental transfers, the Russian government pursued substantial reform of the federal-regional transfer system during the early transition, particularly from 1994 onward (Martinez-Vazquez and Boex 2001). However, until recently, less emphasis was placed on the regional-local systems of intergovernmental transfers. As a result, most regional governments continue to distribute shared tax revenues based on "regulated" sharing rates (allowing local governments in greater need to retain a greater share of revenues) and simple "gap filling" transfers (filling the gap between actual expenditures and revenues) to equalize fiscal disparities across their local governments.

This chapter describes the federal-regional transfer system and regional-local transfers including shared tax revenue, describes federal policy regarding intraregional fiscal equalization, examines vertical and horizontal fiscal imbalances at the subnational level, looks at how well the system of transfers has worked so far, and suggests options for improving the design and performance of transfers between regional and local governments. Appendix 6.1 provides a primer on each of the different steps in the development of a formula-driven system of *horizontal* equalization. As we have been arguing in this book, and it is a fact well known to fiscal decentralization experts, the only complete and also flexible way to reach *vertical* balance is to provide all levels of government with sufficient revenue autonomy. However, in the context of Russia's budget institutions, uniform and stable sharing of tax revenue on a derivation basis is a clear second-best policy in comparison to annually budgeted grants. We return to this topic in the discussion of the Kozak Commission proposals toward the end of this chapter and also in appendix 6.2 surveying some of the international experiences with fraternal or horizontally funded transfers.

The Federal Transfer System and Regional-Local Fiscal Transfers

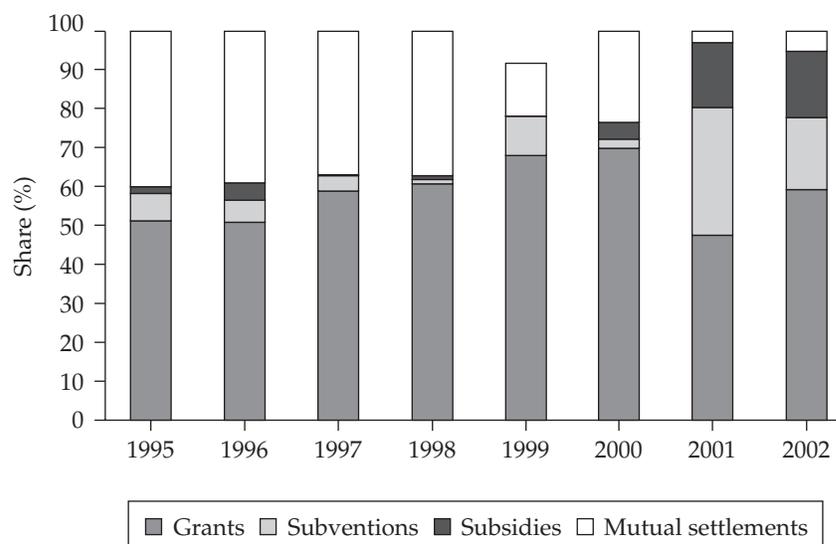
An appropriate starting point for the discussion on regional-local transfers in the Russian Federation is the country's federal-regional transfer system. Not only does this system determine the level and scope of resources available for subsequent regional-local distribution, but it also serves as an important model for formulating regional-local transfer systems. In 1994, rates for sharing tax revenue between the federal government and regional governments were in effect standardized across regions (while varying by tax) and since that year have been annually set

in the federal budget law virtually unchanged.¹ The implication was that all regions, be they rich or poor, had to be treated the same. Therefore, the attainment of more efficient and transparent revenue assignments brought about potentially larger fiscal disparities in the Russian Federation. The federal government had to rely on a new system of equalization grants to offset some of these increased disparities in the distribution of fiscal resources.

Figure 6.1 shows how the total amount of federal transfers to Russia's regions is distributed under four major categories: equalization grants, subventions, subsidies, and mutual settlements. The most dominant federal-regional grant modality is the equalization grant (the Federal Fund for the Financial Support of the Regions). The size of the federal equalization grant pool is set prior to the beginning of the fiscal year in the annual federal budget law. Although the shares of individual regions in the equalization fund are drafted based on the need/capacity formula, the legislated amounts may end up considerably different as a result of political bargaining in Parliament, in particular over specific weights to be used in the formula.²

Moreover, actual shares of regions in the equalization fund may differ from the legislated ones if the total size of the fund falls below the planned amount and the Ministry of Finance adjusts the planned allocation unevenly across regions. As can be seen from figure 6.1, the share

Figure 6.1. *The Composition of Federal Transfers to Regional Governments*



Note: 1999 figures do not add up to 100 percent owing to the presence of unidentified "other" transfers.

Source: Authors.

1. Whereas revenue-sharing arrangements had been relatively stable from 1994 through 1997, several substantial changes were made in revenue-sharing arrangements before and after the economic crisis of August 1998 in the direction of recentralizing tax revenues. Changes in the assignment of revenues, especially the assignment of PIT collections, reflected a tug of war over resources between the federal and regional governments. As shown in table 6.1, fiscal policy reforms in recent years have aimed at reassigning tax revenues away from subnational governments and toward the federal government. In addition to the centralization of PIT revenues, the reassignment of the VAT to the central government most significantly contributed to the centralization of public-sector resources. The federal government's share of overall tax collections reached 60 percent in 2001 from a low of 42.5 percent in 1997.

2. The transfer formula undergoes some changes every year. Thus, for the 2005 transfers, the Index of Budget Expenditures will be computed for the composite subnational expenditures rather than 15 separate expenditure categories as in previous years. However, the set of factors used in the formula is quite stable and traditionally has included variables accounting for differences in wage costs, price level, utility costs, accessibility of the region's localities by transport, and demographic composition of the population.

of equalization grants in total federal transfers increased from 50 percent in 1995 to 70 percent in 2000. The remaining share mostly goes to “mutual settlements,” which are typically not budgeted *ex ante*, but rather result *ex post* from emergency situations and political lobbying. The relative decline of mutual settlements over time is an indication of improvements in the objectivity, stability, and predictability of the federal-regional transfer system. In addition, regional governments have been provided with intergovernmental loans for liquidity purposes. These loans can be considered additional grants because they are often rescheduled and eventually offset or written off.

Since 2001, equalization grants, which are unconditional, have been complemented with earmarked transfers to reimburse subnational governments for the two major federal mandates: payment of monthly child benefits (reported in the budget as “subventions”) and subsidies for payment on various goods and services by persons with disabilities (reported in the budget as “subsidies”).³ These earmarked grants work as partial replacement of the equalization grant, as the latter used to account for the mandates’ burden in the assessment of regions’ need. At the same time as these two large conditional grant programs were introduced, the regional share of VAT collections was entirely centralized to the federal budget for subsequent redistribution across regions via the earmarked grants.

The centralization of VAT collections resulted in a significant increase in regional governments’ dependence on federal transfers in 2001. This dependence can be measured as a share of subnational expenditures financed by federal transfers. As shown in table 6.1, there are wild variations across regions in this aspect of revenue autonomy; while some regions receive little or no federal transfers whatsoever, other regions are almost completely dependent on them. As can also be seen from the table, the average dependence of the consolidated regional-local government on federal transfers has been increasing consistently from 25 percent in 1999, reaching 35 percent in 2002. However, at the same time the extent of variation among regions of dependency on federal transfers has been narrowing, as signaled in the decline of the coefficient of variation over the same period. Therefore the outcome of the recentralization appears to be a more uniform share of the subnational expenditures’ burden being eased with federal transfers.

At the regional-local level there has been much less progress. As already pointed out above, most regional governments continue to use “regulation” for tax sharing and simple “gap filling” transfers to equalize fiscal disparities across their local governments. Although there have been some reform efforts by a group of regions, most Russian regional governments continue

Table 6.1 *Regional-Local Dependence on Federal Transfers (percent)*

	1995	1996	1997	1998	1999	2000	2001	2002
Mean	22	26	23	26	25	27	33	35
Median	14	23	22	21	17	20	29	30
Minimum	0	0	0	0	0	0	1	1
Maximum	81	68	66	93	98	92	90	87
Coefficient of variation	0.92	0.64	0.68	0.86	0.93	0.90	0.70	0.66

Note: Federal transfers are shown as a percentage of consolidated (regional-local) expenditures.

Source: Calculated from Russian Ministry of Finance data.

3. Until 1998, the budget item on “subventions” reported compensation to the City of Moscow for the costs related to its status of the nation’s capital. In 1999, the “subventions” item covered the federal government’s aid to regions in replenishing “supplies of necessities” in localities of the Far North. Being a relic of the Soviet planners’ decision on location, the northern settlements still have to be subsidized on humanitarian grounds before an eventual relocation is carried out.

to employ budgetary practices inherited from the Soviet system. It appears that in many regions, public servants continue to think and act as they did under the previous regime. For example, despite being given greater autonomy than under the previous system, subnational governments frequently do not take the initiative and instead wait for guidelines from the higher authority. In particular, the function of finance officers used to be merely to aggregate budget requests from constituent lower-level governments and public entities and pass them on for funding to the higher authority. However, when the federal authorities stopped considering actual costs and switched to formal indicators of expenditure needs, regional governments had to revise their relations with localities.

It was only in the late 1990s that federal legislation attempted to bring some structure to the regional-local fiscal relations. The federal *Law on the Financial Foundations of Local Self-Government in the Russian Federation* (1997) prescribed that subjects of the Federation should establish a “Fund for the Financial Support of Municipal Settlements” with the following characteristics:

- “The means of the fund . . . shall be amassed by way of deductions from federal and regional taxes coming into the budget of the subject of the Russian Federation.”
- “The means . . . shall be distributed according to a fixed formula, which would take into account the municipal settlement’s population, the share of preschool and school children in the municipal settlement’s population, the share of pension-age persons in the municipal settlement’s population, the area of the municipal settlement’s territory, and the level of the per capita provision with the budgetary means in the municipal settlement, as well as certain other factors accounting for the specifics of the given subject of the Russian Federation.”
- “The share of every municipal settlement in the total volume of the means in the fund for the financial support to the municipal entities shall be fixed in percentages and shall be approved by the law of the subject of the Russian Federation on the budget of the subject of the Russian Federation.”

This piece of federal legislation attempted also to secure the financial independence of local governments by requiring that for major taxes, regional governments should, on average, meet certain minimum rates of sharing with local governments. This was considered overall a toothless provision because it included the phrase “on average” and the vast majority of regions had been already in compliance. The same law also required regions to set and secure minimum expenditure standards for local governments.

Although many regions responded to the federal requirements by establishing an equalization fund with the name prescribed in federal legislation, the distribution of these funds’ grants hardly follows the letter of the federal law (see box 6.1 for a description of the system in Tomsk Oblast, which functions in accordance with federal requirements). While implementation varies by region, generally it merely “repackages” the Soviet gap-filling approach. Only a few regions have introduced a formula for grant distribution while others still resort to annual negotiations on the basis of historical tax collections and expenditures.⁴ Even in regions using a formula, the employed expenditure norms are mostly based on the existing facilities’ capacity rather than service need (Kourliandskaia, Nikolayenko, and Golovanova 2002). As a rule, facility-driven expenditure needs estimated with bottom-up-derived norms exceed the available resources thus resulting in a fiscal gap. Often the distribution of regional grants aims at leveling this fiscal gap among localities. Although essentially this is equivalent to downscaling the initial expenditure norms to match the available resources, this back-and-forth process hampers the transparency of the grant system.

4. To our knowledge the regions that introduced equalization formulas include Leningrad, Sverdlovsk, and Moscow oblasts. For details on the Leningrad oblast formula see Bahl and the Intergovernmental Fiscal Relations Team (1999).

Box 6.1 *Distribution of Equalization Grants in Tomsk Oblast*

The procedures for grant distribution in Tomsk Oblast are described in the document titled *Intergovernmental Relations in Tomsk Oblast*, so that the process is rather transparent unlike in many other regions. The calculation of grants starts with estimating the “minimum local budgets.” Minimum local budgets are derived in different ways for each category of expenditures, although virtually all are based on per unit spending. For every category the average previous-year spending per capita in the oblast, adjusted for expected inflation, is multiplied by the number of units in the locality. For some categories, the units’ multiplier is driven by the number of “customers” (for education, numbers are broken down into pre-school, day school, boarding school, orphan homes, and after-school center client-groups), but sometimes the customer group is simply defined as the population of the locality. And for other categories the needs-driver is infrastructure (square meters of housing) or norm-based (for example, required number of municipal employees, norm-based number of libraries, club seats, or museums).

The calculation of “minimum local budgets” in Tomsk does not establish the amounts that localities are expected to spend; it is only used for the distribution formula. The transfer amount for each locality is determined as minimum expenditure requirements less revenue forecast for this locality. The amount of the oblast transfer fund is determined by the total of the expected fiscal deficits across localities. As a result, a sizable percentage of oblast tax revenue is dedicated to the fund for distribution to the localities.

The equalization outcome suggests that assistance is indeed directed to those localities with lower economic capacity and greater expenditure need. The grants are greater for localities with a smaller economic base and they do respond to the numbers of the population more likely to need local services (pensioners and those under working age). The number of housing units, clubs, libraries, and museums do not have strong influences on the amount of the grant received, but there is a positive relation with the number of schools (in other words, the presence of more schools in per capita terms means greater grant assistance).

Source: “Tomsk Oblast Assessment Report,” Georgia State University/USAID Fiscal Reform in the Russian Federation Project (1999).

Federal Guidance on Intraregional Equalization

Theoretically, there are four main types of federal policy on regional-local fiscal relations:

- The federal government could strictly impose a particular system of regional-local fiscal relations (the imposed system could be uniform across regions or vary by type of region).
- The federal government could offer regions a choice of several regional-local fiscal patterns, which are nevertheless structured at some level of detail.
- The federal government could limit itself to providing general guidelines and structure that must be followed by all regions.
- The federal government could give each region freedom to define its own relationship with constituent localities.

The last two options would seem to be most appropriate for a federal country, but given the nature of Russian federalism, the first two options cannot be excluded from consideration.

In actuality, the federal government’s current policy-design approach toward regional-local fiscal relations presents a mix of the above four alternatives. In areas where regional governments have some degree of discretion, they may need some technical assistance on how to make their own choices. For example, in the case of regional-local transfers, the federal government could suggest that the new federal-regional model be applied to the regional-local level. However, in a number of aspects the regional-local sector differs from the federal-regional, and also exhibits significant variation across regions. For example, on average, decentralization of expenditures is higher in the regional-local sector than in the federal-regional sector. At the same time, the share of own and assigned revenue is larger in regional than in local budgets. Dependence of many regions on federal grants further aggravates vertical imbalances within regions. Moreover, unlike regions, which are constitutionally all equal subjects of the Russian

Federation, local governments are not all equal and actually differ by type: urban versus rural and territorial versus settlement-based.

The lack of widespread reform initiatives at the regional-local level has been continuing to put pressure on the federal authorities to act. In July 1998, the federal government approved a program of reforms titled *Concept of Restructuring Intergovernmental Fiscal Relations in the Russian Federation for 1999–2001*.⁵ Besides important changes at the federal-regional level, the program also concerned regional-local fiscal relations, as reflected in items 21 and 22 of the implementation plan. These items provided for the development of methodological recommendations to regional governments on the norm-based assessment of local expenditure needs as well as the assessment of local fiscal capacity and allocation of financial assistance to localities.

The first piece of federal guidance on intraregional equalization was delivered in the *Interim Methodological Recommendations to RF Subjects on Arranging Intergovernmental Fiscal Relations* issued by the Russian Ministry of Finance in June 2000. The *Recommendations*, focusing on regional-local relations, resulted from several months of discussions in the standing Intergovernmental Fiscal Relations Task Force at the Ministry of Finance. The task force was composed of representatives of the federal Ministry of Finance, finance officials of selected regions, and invited experts. In addition, the task force received some input from nonrepresented regions in the form of comments on drafted materials. The *Recommendations* include the principal text, explanatory notes, and Excel-based simulations for three Russian regions: Vladimir, Karelia, and Rostov. The main text provides an assessment of intergovernmental fiscal-relations arrangements prevailing within Russian regions; outlines the principles of optimal intergovernmental fiscal-relations arrangements based on the *Budget Code* provisions and best practice; and offers a concrete mechanism for allocating regional transfers to localities. The explanatory notes provide more details on the analysis of the status quo, elaborate on intergovernmental fiscal-relations principles, and discuss modalities for the implementation of the transfer formula.

The development of the *Recommendations* marked a significant step forward in the integral reform of the Russian system of intergovernmental fiscal relations. Also, the reliance on the task force created a precedent for the exercise of joint authority of the federal and regional governments over local governance issues, as prescribed by the 1993 Constitution. In addition, the task force discussions enabled horizontal exchange of regional experiences, as several participants represented pilot regions with internationally assisted reform projects.⁶ The development of the *Recommendations* was very timely and relevant for regional governments, which required some technical assistance in implementing the new principles of budgeting introduced with the *Budget Code* and in devising their relationship with constituent municipalities in particular. It must be noted that the featured transfer formula was an adaptation of the mechanism used for the distribution of federal grants to regions. This approach appealed to many regional governments, which, in their relationship with constituent localities, often attempt to mimic the federal government's relationship with the regions.

The Extent of Fiscal Equalization at the Subnational Level

The large disparity in own-source and assigned revenues among jurisdictions within one region inevitably means that in the absence of an equalization mechanism some localities

5. Resolution No. 862 of the RF Government of July 30, 1998. The major focus of the program was on the methodology for the allocation of the Federal Fund for the Financial Support of the Regions (FFSR), that is, the federal-regional equalization grant system. The main aim was to transform indicators of expenditure needs and revenue capacity from actual levels to some objective assessments by 2000. Another important component of the program was to inventory federal regulations affecting subnational expenditures, including unfounded federal mandates.

6. These were regions participating in the Fiscal Reform Project implemented by Georgia State University, funded by USAID (1997–2000).

would not be able to spend roughly similar amounts per capita in the provision of services to their residents. Not surprisingly, many regional governments have continued to rely on the “regulation” of taxes as the main methodology for closing horizontal fiscal imbalances among local governments. The approach consists of lowering revenue-sharing rates for one or more taxes in richer regions and increasing them in poorer regions. This approach was also used by the federal government between 1992 and 1994. The practice of relying on “regulated” tax-sharing rates despite all the problems it represents (involving bargaining, discretion, uncertainty, instability of funds, and negative incentives to revenue mobilization) continues to present significant advantages to regional authorities for equalizing fiscal disparities across jurisdictions, as the mechanism is simple and readily understood by local government officials.⁷

The desired target level of equalization is fundamentally a policy choice, albeit one that should be informed by technical considerations. While in many regions the allocation of regulated tax revenues and grants by the regional authorities significantly reduces the initial fiscal disparities, this approach to equalization also means that localities raising more revenue—either through stronger economic activity or greater tax effort—would not be able to provide better services for their residents. Regional policymakers thus have to balance the benefits of greater equity in local service provision against the losses in public revenue and economic activity caused by the potential negative fiscal incentives imbedded in the equalization scheme.

In reality, the allocation of regional-local intergovernmental grants may also be driven by political goals unrelated to fiscal equalization. For example, a regional government might use its control over equalization grants to reward local allies or punish local authorities representing a rival political party. This is not an unlikely scenario for Russia, since often regional governments come to power with the support of the rural vote and thus have political agendas rival to that of urban mayors. Similar political influences over federal-regional allocations earlier in the transition were uncovered by Treisman (1996) and Stewart (1997). As a result, the political choice for the desired level of equalization is hard to trace from formal sources. Instead, the desired degree of equalization is implicitly revealed by the local fiscal disparities that exist before and after equalization.

How Big Are the Fiscal Disparities at the Subnational Level?

Table 6.2 shows the differences in per capita revenue among local governments (before and after cumulatively adding assigned revenues, regulated revenues, and intergovernmental transfers) for local governments in the 73 regions (2,038 localities total) of the Russian Federation for 2001.⁸ In that year, the per capita amount of own-source revenues plus assigned revenue ranged from zero rubles in Sunzhensk Rayon (Ingush Republic) to 56,633 rubles per person in Yak-Bodinsk Rayon (Udmurt Republic), with the median amount of 529 rubles per person. The coefficient of variation in own-source and assigned revenue stands at a rather high level of 2.61. This disparity is somewhat lessened after the allocation of regulated tax revenues among localities, so that the coefficient of variation decreases from 2.61 to 2.36. However, considerable equalization is achieved only after the distribution of regional-local grants, after which the coefficient of variation drops sharply to 1.53.

7. At the federal-regional level, the switch to uniform sharing rates in 1994 implied that all regions, rich and poor, had to be treated the same. Therefore, the attainment of more efficient and transparent revenue assignments brought about potentially larger fiscal disparities in the Russian Federation. As a result, the federal government’s introduction of uniform sharing rates coincided with the introduction of a new system of equalization grants to offset some of these increased disparities in the distribution of fiscal resources. In addition, the federal government maintained an ad hoc mechanism of “mutual settlements.” Finally, the federal government further provided regional governments with intergovernmental loans for liquidity purposes (more on this in chapter 8). These loans can be considered additional grants because they are often rescheduled and eventually offset or written off.

8. See also footnote 4 in chapter 4.

Table 6.2 *Variation in Per Capita Revenue of Local Governments within and between Regions, 2001*

	<i>Own-source</i>	<i>Plus assigned</i>	<i>Plus regulated</i>	<i>Plus grants</i>
Grand mean, RUR	727	1,226	2,423	4,923
Grand median, RUR	337	529	1,269	3,389
Grand maximum, RUR	32,968	56,633	120,566	129,081
Grand minimum, RUR	-137	0	4	17
Median of regional means, RUR	500	819	1,773	3,895
C.V. (total)	2.41	2.61	2.36	1.53
C.V. between regional means	1.96	2.80	2.55	1.80

C.V. Coefficient of variation.

Note: Negative amounts stem from refunds for past collections.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

The true extent of cross-jurisdictional equalization achieved with the allocation of regulated tax revenue and grants is somewhat disguised in the descriptive statistics presented in table 6.2—because we pooled localities from different regions. The last row of table 6.2 shows that the coefficient of variation between regional means of local revenue is even larger than for the coefficient of variation for localities pooled across different regions (except for the initial own-source revenue). Table 6.3 reports coefficients of variation for per capita local resource allocations within individual regions for cumulative steps in the local resource allocation process. Intraregional variations in revenue collections are substantially smaller than inter-regional variations; for own-source and assigned revenues the median coefficient of variation within regions is three times smaller than the overall coefficient of variation computed for the pooled sample. Moreover, with the regulated tax revenue and grants included, the median coefficient of variation within regions becomes smaller than the coefficient for the pooled sample by a factor of 4.5. As a result of equalization, in a median region per capita fiscal resources of local governments differ by no more than 40 percent, while for the pooled sample this variation is 150 percent.

Similar to the trend revealed by the pooled (inter-regional) set, consideration of the mean and median values in table 6.3 suggests that each consecutive step in resource assignment results in a more equal allocation of local resources within each region. However, looking at the median extent of equalization in table 6.3 still does not show us the entire picture. In fact, comparing the coefficients of variation before and after “equalization” (effected through regulated tax revenues and grants) for individual regions shows a dramatic decrease in local fiscal disparities in some regions while almost no equalization is achieved in others. Moreover, further analysis (not shown in the table) suggests that the initial level of inequality does not have much bearing on the extent of equalization finally undertaken within a region. When the sample is divided into “heavily equalizing” and “hardly equalizing” regions, for both groups the initial median coefficient variation is roughly equal to the total sample median. More rigorous evidence is provided by the fact that the correlation coefficient between the initial coefficient of variation and the after/before ratio is merely 0.07.

It is also interesting to track how the extent of disparity is transformed after each cumulative stage of revenue allocation by observing the coefficient of variation for own-source revenues, and then noting changes in the coefficient of variation after consecutively adding assigned tax revenues, regulated revenue, and intergovernmental grants. Assigned tax revenues (from corporate assets, retail sales, and natural resources) slightly lessen the fiscal disparity between localities within regions (table 6.2) but considerably increase the disparity between regions (table 6.3). The latter effect is quite intuitive given the uneven distribution of natural resources (and to a lesser extent, corporate assets) across regions, which also affects the retail sales indirectly through the income effect on consumption and retail. However, the equalizing effect (albeit only very small) of the assigned revenue on the within-region disparities is

Table 6.3 *Coefficients of Variation of Local Revenues in 73 Regions, 2001*

<i>Oblast code</i>	<i>Subject of RF</i>	<i>Own-source</i>	<i>+ Assigned</i>	<i>+ Regulated</i>	<i>+ Grants</i>	<i>After Grants/ After Assigned</i>
1	Republic of Bashkortostan	1.07	0.79	0.71	0.32	0.41
2	Republic of Buriatia	1.44	1.55	1.32	0.55	0.36
3	Republic of Dagestan	2.31	2.06	1.49	0.38	0.18
4	Kabarda-Balkar Republic	0.61	0.62	1.51	0.42	0.68
5	Republic of Kalmykia	1.72	2.09	2.17	1.32	0.63
6	Republic of Karelia	1.60	1.39	1.08	0.72	0.52
7	Republic of Komi	1.00	1.32	0.62	0.22	0.16
8	Republic of Mari El	0.73	0.69	0.86	0.16	0.23
9	Republic of Mordovia	0.58	0.58	0.49	0.25	0.43
10	Republic of North Osetia	0.94	0.73	1.26	0.44	0.60
11	Republic of Tatarstan	1.85	1.69	1.48	0.84	0.50
12	Republic of Tyva	1.38	1.49	1.73	0.59	0.40
13	Udmurt Republic	3.78	3.10	3.45	2.79	0.90
14	Ingush Republic	1.13	1.41	1.18	1.34	0.95
15	Chuvash Republic	0.87	0.72	0.71	0.14	0.20
18	Krasnodar Krai	0.69	0.71	0.61	0.40	0.56
19	Krasnoyarsk Krai	2.35	2.31	1.68	0.81	0.35
20	Primorski Krai	0.91	0.76	0.63	0.19	0.25
22	Khabarovsk Krai	1.11	2.21	1.29	0.76	0.35
23	Amur Oblast	1.39	1.22	0.69	0.46	0.37
24	Arkhangelsk Oblast	0.74	0.73	0.46	0.19	0.26
25	Astrakhan Oblast	1.43	1.61	0.86	0.34	0.21
27	Bryansk Oblast	0.87	0.76	0.76	0.35	0.45
28	Vladimir Oblast	0.57	0.55	0.51	0.26	0.48
29	Volgograd Oblast	0.65	0.90	0.68	0.17	0.19
30	Vologda Oblast	0.70	0.66	0.54	0.15	0.23
31	Voronezh Oblast	0.80	0.76	0.69	0.24	0.32
32	Nizhny Novgorod Oblast	1.03	0.94	0.74	0.26	0.28
34	Irkutsk Oblast	0.75	0.91	0.64	0.58	0.63
35	Kaliningrad Oblast	1.32	1.65	1.39	0.77	0.47
36	Tver Oblast	0.90	0.82	0.68	1.00	1.21
37	Kaluga Oblast	0.74	0.73	0.72	0.25	0.34
38	Kamchatka Oblast	0.93	0.83	0.52	0.47	0.57
39	Kemerovo Oblast	2.92	2.61	2.54	1.95	0.74
40	Kirov Oblast	0.83	0.63	0.42	0.14	0.22
41	Kostroma Oblast	2.90	2.69	2.52	1.70	0.63
42	Samara Oblast	0.68	0.58	0.87	0.51	0.89
43	Kurgan Oblast	0.42	0.48	0.56	0.08	0.17
45	Leningrad Oblast	0.97	1.01	0.75	0.43	0.43
46	Lipetsk Oblast	0.63	0.54	0.50	0.17	0.31
47	Magadan Oblast	0.41	1.00	0.68	0.39	0.39
48	Moscow Oblast	0.74	0.67	0.44	0.29	0.43
49	Murmansk Oblast	0.51	0.51	0.34	0.25	0.50
50	Novgorod Oblast	1.47	1.61	1.48	1.15	0.71
51	Novosibirsk Oblast	0.60	0.70	0.61	0.23	0.33
52	Omsk Oblast	0.71	0.75	0.76	0.30	0.39
53	Orenburg Oblast	0.74	0.88	0.62	0.38	0.43
54	Oryol Oblast	0.96	0.74	0.58	0.15	0.21
55	Penza Oblast	0.57	0.57	0.50	0.16	0.27
56	Perm Oblast	1.13	1.24	0.73	0.29	0.24
57	Pskov Oblast	0.69	0.59	0.48	0.20	0.34
58	Rostov Oblast	0.60	0.53	0.52	0.18	0.33

(Continues)

Table 6.3 *Coefficients of Variation of Local Revenues in 73 Regions, 2001 (Continued)*

<i>Oblast code</i>	<i>Subject of RF</i>	<i>Own-source</i>	<i>+ Assigned</i>	<i>+ Regulated</i>	<i>+ Grants</i>	<i>After Grants/ After Assigned</i>
59	Ryazan Oblast	1.37	1.25	0.71	0.22	0.18
60	Saratov Oblast	0.77	0.70	0.60	0.26	0.37
61	Sakhalin Oblast	1.32	1.00	0.73	0.54	0.54
62	Sverdlovsk Oblast	0.78	0.68	0.58	0.32	0.47
63	Smolensk Oblast	0.67	0.90	0.82	0.36	0.40
64	Tambov Oblast	0.55	0.53	0.52	0.15	0.29
65	Tomsk Oblast	1.46	1.56	1.24	0.47	0.30
66	Tula Oblast	0.74	0.73	0.63	0.23	0.31
67	Tumen Oblast	1.24	1.19	0.98	0.39	0.33
68	Ulianovsk Oblast	0.94	0.75	0.69	0.28	0.37
70	Chita Oblast	2.12	1.47	1.59	0.68	0.46
72	City of St. Petersburg	0.98	0.85	0.79	0.19	0.23
76	Republic of Adygeya	0.98	0.84	0.57	0.23	0.27
78	Jewish AO	1.06	0.93	0.83	0.79	0.85
79	Karachayevo-Circassian Republic	1.13	1.50	1.28	0.47	0.31
80	Republic of Khakasia	0.76	0.71	0.66	0.20	0.28
82	Komi-Perm AO	0.63	1.07	0.90	0.37	0.35
83	Koryak AO	0.56	0.55	0.53	0.49	0.88
87	Khanty-Mansi AO	0.59	1.03	0.73	0.43	0.42
89	Evenk AO	0.90	1.06	0.71	0.33	0.31
90	Yamal-Nenets AO	0.74	0.88	0.62	0.38	0.43
	Mean	1.07	1.07	0.92	0.48	0.42
	Median	0.90	0.84	0.71	0.34	0.37
	Minimum	0.41	0.48	0.34	0.08	0.16
	Maximum	3.78	3.10	3.45	2.79	1.21

AO. Autonomous oblast.

Source: Authors' calculation based on data from the Center for Fiscal Policy, Moscow.

hard to explain, as the aforementioned taxable bases are not likely to be evenly distributed among localities of one region.⁹ One possible explanation is that rural areas might have more retail sales than general sales, which would dominate the own-source revenue. Another possibility is that some unaccounted regulation might be going on through the allocation of the corporate assets tax in excess of the legislated 50 percent retention rate.

The cumulative allocation of regulated tax revenues to local governments (including the VAT, CIT, PIT, and excises) slightly lessens the local fiscal disparities between regions (table 6.2) and achieves a bit more equalization within regions (table 6.3). Both come as a surprise: why is equalization achieved at all in the first instance, and why so little in the latter? The consolidated regional-local share of regulated taxes should essentially act as an assigned-revenue source because of the uniform remittance rates to the federal level. Thus, there are no clear reasons why regulated taxes should bring about any equalization between regions. One possibility is that regional governments in poorer regions allow local governments to retain a larger share of the regional-local tax collections since these are negligible compared to the amount of federal grants received by such regions (Timofeev 2003).

9. Another puzzle with the assigned revenue is that its share in total local government revenue is positively related to the vote gained by the Communist candidate in the 1996 presidential election. As the Communist Party's performance is likely to be related to the sentiment toward centralist/redistributive policies, there might be a (so far unexplained) link between the local yield from assigned revenue and the level of equalization within Russian regions.

In contrast, the weak within-region equalization effect of regulated tax revenue might be explained by the fact that in many regions this source of revenue might, in effect, be treated more like an assigned tax than an intergovernmental transfer or grant. Furthermore, it must be noted that even though these revenues still receive the adjective “regulated,” there may be little left here of the original meaning of “regulation,” where tax rates were scaled up or down inversely to the relative wealth of the local jurisdictions. The lack of data on total collections of regulated tax revenue in a local territory precludes us from exploring some obvious interesting questions such as what the extent of variation in sharing rates is for major taxes, or what determines these sharing rates, and whether these sharing rates are negatively related to the size of the tax base or tax effort. One reason why regions might be treating regulated taxes more or less like assigned taxes is that federal legislation requires regions to meet some minimum sharing rates for this type of revenue with local governments, even if only on average. In addition, while nominal amounts of grants become eroded with annual inflation, the last year’s retention rates present a more stable benchmark for a local government bargaining with regional authorities. Nevertheless, as was the case with our observation for the pooled sample, we see that most of the within-region equalization occurs only through transfers or grants.

The Need to Focus on Important Issues

In many regions, the current regional-local transfer systems used by the regional government fail to follow sound principles in the design of intergovernmental transfer systems. Regional-local transfer systems are generally inefficient because they tend to ignore important issues concerning the local tax effort and fail to emphasize efficiency in public expenditure choice. The transfer mechanisms in use at the regional level often still embody undesirable features including too much negotiation and bureaucratic discretion in the determination of the size and distribution of transfer funds (rather than relying on a truly formula-based system), the introduction of negative incentives represented by the claw-back of any “excess” local revenues, the frequent use of confusing conditionalities attached to the equalization grants, and the absence of a complementary transfer mechanism based on special purpose grants.

Negotiation and discretion. Despite the principles set forth in the *Law on the Financial Foundations of Local Self-Government in the Russian Federation* of 1997, in many regions the system has continued to work in a discretionary manner, pretty much along the traditional line of negotiating a “gap filling” transfer for each local government that closes the gap between a locality’s negotiated level of “expenditure needs” and actual local revenues. This practice has continued to detract from the transparency and objectivity of the system of transfers. It should be emphasized that the problem does not necessarily lie in the amount of funds available for transfers, but rather in the manner in which this amount is adjusted and distributed. Admittedly, the lack of adequate data, as well as macroeconomic and overall budget balance conditions, may demand sporadic adjustments to the level of funding and the distribution formula. Nevertheless, the entire process of allocation—that is, the determination of the overall pool of funds and how the funds will be distributed—should rest much more on rules.

Lingering negative incentives in many regions. There is empirical evidence that regional governments offset local governments’ increasing own-source revenue with decreasing grants. Zhuravskaya’s (2000) data set covers 35 large Russian cities in 29 regions, while Alexeev and Kurlyandskaya’s (2003) study focuses on poorer municipalities concentrated in one region. Neither study can reject the hypothesis that the regional governments offset changes in municipal revenues with changes in subsidies. However, Alexeev and Kurlyandskaya show that this offsetting is likely to take place through the adjustment of next year’s transfers, rather than through the current year’s.

Conditionality in the implementation of equalization transfers. If equalization transfers could be correctly conceived as unconditional (general funding) grants, their final use should conceptually

be left to the discretion of local governments. However, in practice, this principle has not been respected or its exercise by local governments has led to political problems. There are many examples of how this happens. The most common practice is for regions to impose priorities on local governments to control the use of funds received from equalization grants. One main problem is that regional authorities often want to pursue multiple objectives with the current equalization grants and thus may require local governments to spend the same ruble on several different regional priorities at the local level.

Absence of other transfer instruments. As a principle of sound transfer design, it would be more effective to use the equalization grants exclusively for equalizing fiscal disparities across local governments, and to make use of conditional transfers to pursue other objectives. Unfortunately, a significant feature at the regional level is the almost complete dominance played by the gap-filling (equalization) transfers and the virtual nonexistence of other types of transfers, such as conditional transfers to pursue regional (for example, sectoral) policy objectives.

The Kozak Commission's Proposed Changes to Intergovernmental Transfers at the Subnational Level

The Kozak Commission has numerous recommendations concerning the improved functioning of regional-local grants. Among others, the Commission's report addresses revenue sharing and "regulation" of taxes (which are intimately related to the effectiveness of regional-local equalization grants), addresses many issues in the creation of equalization grants per se, and recommends the introduction of special local compensatory grants for unfunded federal and regional mandates on local governments.

An efficient and incentive-compatible system of equalization grants requires revenue assignments (and particularly, revenue-sharing arrangements) that are transparent and stable. To this end, the Kozak proposals would extend the principle of uniformity of tax-sharing rates from the regional level to the local level. In particular, the sharing of tax revenues within a region would only be allowed if all jurisdictions of the same tier of government are entitled to the same retention rate. This "entitlement" can be further reassigned by the receiving (rayon-level) government to the tier below (settlements), but again only if this is done uniformly for all jurisdictions in that tier. Following the Kozak proposals would mean that these entitlements and re-entitlements for local governments could not be limited in time. The passage of these proposals would represent pretty much the end of the concept of "regulation" as has been used by regional governments since the Soviet times.

Appropriately, the Kozak Commission recommends the creation of an equalization fund at the regional level to address the disparities in fiscal capacities among local governments resulting, among other things, from the uniform sharing of tax revenues. The proposals state that equalization at the local level should be performed on the basis of population size of the recipient government and its (normalized) fiscal capacity. At any tier of government (rayon or settlement), transfers would go only to those jurisdictions with per capita fiscal capacity below a particular threshold set for that tier. Quite rightly, the Kozak proposals prohibit basing the measure of fiscal capacity on the actual or planned figures of revenues or expenditures of individual government recipients. The proposals stop short of asking for any further requirements for an objective distribution of the equalization funds, such as using an explicit formula. Instead, the Kozak proposals refer to the *Budget Code*, which currently only requires the fulfillment of "social standards" in distributing transfers.¹⁰ Despite the abolition of regulated (nonuniform) revenue sharing, the Kozak proposals would allow for the allocation of local equalization grants through tax-revenue retention.

10. In the current state of the federal legislation, the notion of "social standards" is poorly defined. It allows for further development either as a financial norm or a service standard.

An important aspect of the Kozak proposals on transfers is that regional governments are envisioned as performing separate equalization processes for districts and for settlements. This means that regional governments would have two separate equalization funds. On the other hand, regional governments would be able to delegate to any district the responsibility for equalizing fiscal disparities among the settlements within that district. In that case the “rayon agent” would receive a subvention from the regional fund of support to settlements (or alternatively a revenue entitlement) to engage in intra-rayon equalization. It is not clear from the proposals whether this delegated equalization of settlements by the rayon government should be separate from the equalization that districts are authorized to perform on their own at their own expense. However, given that the methodology for this own-source equalization has to be legislated by the regional government, the result is likely to be the same regardless of whether the two equalizations are performed separately or jointly.

Also important in the Kozak proposals is that regions would be authorized to introduce fraternally funded equalization systems (dubbed “Robin Hood”) for their local governments—which means the annual budget law of the regional government would be able to mandate negative transfers from a number of municipal budgets to the regional equalization fund. According to the current proposal, the contributions can be required from those municipalities whose per capita fiscal capacity exceeds the fiscal capacity average for that tier of government by a factor of two or more. The size of the contribution (or the negative transfer) is limited to half of the excess over double the average capacity. If not offered voluntarily by corresponding local governments, these contributions can be enforced by regional governments through withholding revenue sharing or even “centralizing” some of the municipality’s own-source revenue. The technical capacity to undertake this “centralization” is not clear given that the Kozak proposals allow localities to have independent treasuries. It is more likely that the donor municipalities would contribute voluntarily under the threat of dismissal for violating the regional law. Equalization systems funded with negative transfers from richer jurisdictions can be flexible and quite effective but they also can be quite disruptive.¹¹ The pros and cons and their relevance to Russia’s subnational governance are discussed in appendix 6.2.

Finally, in the area of regional-local transfers the Kozak proposals require that a special compensatory fund be created in the regional budget that provides resources for the financing of federal and regional functions delegated to local governments. Federal and regional budget laws should determine the size of these funds for the coming budget year on an annual basis. The regional fund would draw from subventions from the federal compensatory fund and other regional revenue sources. Subventions from the regional compensatory fund would be distributed to municipalities proportionally to the number local residents or the size of other local client-population groups that determine or drive the amount of resources required for local governments to fulfill the delegated federal or regional functions. This proposal reflects a much-needed move and adequately addresses the very difficult issue of unfunded mandates at the subnational level.

Options for Further Reform of the Transfer System at the Subnational Level

The Kozak proposals go a long way toward rationalizing and modernizing the system of transfers at the subnational level in the Russian Federation. However, there are some issues that will require further attention to get subnational government transfers on the right footing.

11. For example, negative transfers will allow regional authorities to deal with outliers (that is, local governments that are comparatively very rich) resulting from uniform tax sharing. But explicit “fraternal” contributions from the richer municipalities to support expenditures in poorer municipalities may create resentment and lack of solidarity among local governments in the region.

Is There a Need for Further Federal Guidance?

Although the Kozak proposals provide a sound framework for regional-local grants (for instance, by effectively eliminating nonuniform regulated revenue sharing and by prohibiting “gap filling” grants), this framework could be strengthened. A possible improvement to the proposals would be the offering of a broader range of possible arrangements and formulas along with the discussion of merits and drawbacks of each of them on the grounds of the normative considerations. Given the great diversity of local environments, it seems appropriate that the federal policy toward regional-local relations provides the right balance between flexibility and uniformity. Therefore, in addition to the familiar formula of federal transfers, federal authorities could recommend regional governments other alternatives for regional-local relations. Ideally these alternatives should come from successful experiments in Russian regions.

Conceptually, it is important at the regional-local level to distinguish between vertical imbalance and horizontal inequality. In Russia, vertical imbalance is inherent in every region owing to the mismatch between negligible local taxing powers and considerable expenditure responsibilities devolved to the local level. However, vertical imbalance should be treated separately from interjurisdictional equalization. It is possible to address vertical fiscal imbalances through one-time strategic decisions concerning long-term assignment of tax instruments and shares of tax revenue. Subsequent revisions might occur with changes in technology and policy priorities, yet not very frequently. The Kozak proposals would enable regional governments to address vertical fiscal imbalances by providing local governments with permanent entitlements to a uniform portion of the local yield from regional and federal taxes. In contrast, decisions on fiscal equalization and programs funding are usually made recurrently. Being inherently counterequalizing, tax-revenue sharing (on a uniform derivation basis) is not suitable for horizontal equalization.¹²

The proposed option in the Kozak Commission’s proposals for allocating grants through tax-revenue retention (and to a lesser extent, the option for negative transfers) conceals a certain danger. It can effectively undo the uniform revenue sharing and replicate the outcome from the differentiated (“regulated”) sharing rates currently used in many regions. In fact, since 2000 the *Budget Code* has required tax-sharing rates to be stable for a period of no less than three years. Nevertheless, most regions still treat revenue sharing as equalization grants and alter retention rates on an annual basis. To secure vertical balance at the local level, federal legislation may have to require the long-term assignments of revenue shares to cover a certain portion of consolidated local expenditures. This option would be less restrictive than the currently mandated minimum rates for sharing revenue from particular taxes with local governments at the point of collection.

The Two Thrusts of Subnational Government Transfer Reform

There are two basic thrusts of reforms targeting subnational transfers that would complement those in the Kozak proposals.

First, every region should introduce a truly formula-driven system of equalization grants. Such a grant system would comprise two stable and explicit rules, one determining how to establish the pool of equalization funds in the regional budget and another determining how to distribute those funds. This formula would equalize fiscal disparities among local governments on the basis of differences in expenditure needs and fiscal capacity. The fundamental steps the government would need to take to reform the system of transfers are summarized in box 6.2 and the different options on the technical aspects of the equalization formula are discussed in appendix 6.1.

12. Note, however, that revenue sharing based on formulas—for example, on per capita terms as is the case in Germany for the VAT—can be equalizing.

Box 6.2 Fundamental Steps for Reforming Subnational Transfers

The systematic reform of the system of intergovernmental transfers at the regional-local level in Russia will require a series of steps, including the following:

- Making clear that equalization transfers are unconditional grants, and that conditional grants will be used to pursue defined objectives of the regional government, such as financing minimum levels of education and health services.
- Deciding whether lower-level local governments should depend for their transfers on the rayon-level government, or whether the transfers to settlements should be administered directly by the regional authorities.
- Explicitly identifying how much equalization should be pursued. Should differences in expenditure needs and fiscal capacity be closed fully or should the system encourage revenue mobilization and keep some additional resources in faster-growing regions?
- Determining the funding mechanism for conditional grants and equalization transfers separately. The “Regional Equalization Fund” and the funding for the conditional grants can be financed by centralizing (or setting aside) the proceeds (or a percentage) of several major taxes and federal transfers or by introducing negative transfers for better-off local governments as suggested in the Kozak proposals.
- Considering the introduction of rules for the determination of the conditional grants and the equalization pool. This can be done in an objective and stable basis—such as a percentage of regional government tax revenue, or of certain taxes, with this percentage regulated by regional law or computed as some moving average of actual collection of a number of years.
- Using formulas explicitly stated in the regional law for the division of the equalization funds among districts and settlements. Along the lines suggested in the Kozak proposals, equalization formulas should support a fair allocation of resources by providing more to local governments with lower tax capacity and greater fiscal needs. Equalization formulas should be stable over a period of years to promote budget stability, be simple and transparent, be comprehensible to all stakeholders, be impermeable to political manipulation or negotiation in any of their aspects, and provide the right incentives for revenue mobilization by local governments.
- Considering whether to introduce reforms “cold turkey,” or whether to introduce some sort of “hold harmless” provision. This is fundamentally a political choice, the former usually being a harder strategy to sell and sustain. Alternatively, a “hold harmless” provision would phase-in the new transfer system over a number of years to smooth out the transition to the new regime and typically make the changes in regional funding more politically acceptable.

The advantage of a formula-driven equalization grant system is that it will lead to a more objective, fair, and predictable distribution of funds. It will establish a sustainable level for the total amount of transfers to be distributed. The distribution formula itself can take different forms, but of course it should be designed carefully, so that efficiency and equity are guaranteed in the system (again, see appendix 6.1). It is important that the equalization transfer mechanism is seen exclusively as an instrument for *fiscal equalization*, instead of as an instrument to solve all sorts of intergovernmental fiscal issues. Issues related to vertical imbalances in general, sporadic regional fiscal problems, or delegation of specific central government competences to specific subnational governments, should be managed through other (sectoral or targeted) transfer instruments. That is, the regular equalization transfer should always be treated as a *general-purpose grant*, and be made on a *nonconditional basis*, leaving the final use of funds to be decided by the local authorities.

The second thrust of the reforms should be the introduction of conditional grants in support of particular regional interests, such as leading local governments to provide a minimally adequate level of education and health services, as well as in support of addressing externalities at the regional level, such as in the areas of environmental control and sanitation. These conditional grants could be formulated with or without matching provisions, depending on the objective of the grant scheme. The use of matching arrangements, whereby the regional and local governments jointly participate in the financing of the program, are more conducive to the

efficient use of resources because they introduce the right incentives for the local ownership of programs. A well-defined system of conditional grants should be able to address a source of friction in the system of intergovernmental finances in Russia's regions, which is the different priority given to social sector expenditures by regional and local governments. Often, regional authorities wish that local governments had channeled more funds to the education and health sectors. Ultimately this is a conflict between two important objectives: respecting local autonomy and giving (more) expenditure priority to areas of importance for regional governments. Again, the creation of a system of special-purpose conditional grants for health, education—and possibly other areas as well—from the regional to local governments, will provide a compromise between those two objectives.

These conditional grants could be computed on a “capitation basis” (that is, per student and per inhabitant).¹³ The use of per client conditional grants for education and health would offer several other advantages. First, these grants would be inherently equalizing, thus addressing another major goal of the decentralization system. Second, because these would be (sectoral) block grants, local authorities still would be able to exercise budgetary autonomy in terms of expenditure priorities (within education and health) and the most efficient method of service delivery.¹⁴

Appendix 6.1. A Basic Methodology for Developing a System of Equalization Transfers

The development of a formula-driven system of equalization transfers—as pursued in the Russian Federation—requires four steps: measuring expenditure needs, measuring fiscal capacity, combining those two into a measure of fiscal gap and deciding how much of it to equalize and how, and phasing in the new system of transfers.

The steps involved in structuring regional-local equalization transfers are essentially the same as the steps followed to develop the federal-regional equalization grant system. Nonetheless, there are aspects of the equalization process that are driven by specific subnational consideration, particularly with regard to the availability of data. With regional-local equalization in mind, this appendix provides a primer on each of the four steps.

Measuring Expenditure Needs

There are several approaches to measuring expenditure needs. For a subnational government these needs may be defined as the funding necessary to cover all expenditure responsibilities assigned to the region at a standard level of service provision. In practice, there are several options to measure differences in expenditure needs across subnational governments:

One option is to measure from the bottom up, by costing the current expenditure obligations of local governments or costing a standardized basket of subnational government services. However, this approach requires all sorts of information and very explicit procedures outlining how to cost all aspects of the subnational government's expenditure responsibilities. Another

13. The per capita basis could be modified, if needed, by some adjustment coefficient to reflect different costs of provision or needs, but only as long as these adjustments can be made by formula and do not involve negotiation among central and local authorities. Preferably, the difference in needs arising, for example, from different levels or quality of infrastructure, would be addressed by sectoral ministries using separate grants disbursed using a formula or objective criteria.

14. It is important to emphasize that the conditional grants should not be made for economic categories of expenditure, such as wages or salaries. This approach would eliminate the advantages of a decentralized delivery and implementation system and the full centralization of those services would produce the same results. Naturally, for local governments to be able to increase efficiency in the delivery of services, it would also be necessary to reduce, if not eliminate, central government norms and mandates *that interfere with their choices*.

disadvantage it has is that there is no guarantee the expenditure needs so derived are affordable within the overall budget resource envelope. The insufficiency of funds will require a downward adjustment of the computed budgets. This often can become a source of frustration for subnational government officials, if not a cause of voter protest.

A second method relies on historical expenditure patterns. In this regard, expenditure levels would have to be adjusted over time for inflation and possible changes in expenditure responsibilities. But, in any case, it is not guaranteed that even the “adjusted historical expenditure patterns” would properly reflect the “actual” needs, since each jurisdiction may have exerted distinct efforts in terms of expenditure efficiency in the recent past. Moreover, tremendous demographic changes make past expenditures a poor indicator of present service needs.

A third, and more commonly used approach is to estimate some type of index of *relative expenditure need*. Implicitly, this is what is done when a weighted-factor mechanism is used to allocate equalization grants. These indexes attempt to capture, from the simplest to more complex ways, the factors that determine cost differences in delivering a standard package of local government services. These factors include demographic variables reflecting, for example, the special needs of the young and the elderly, other factors such as the level of poverty and unemployment, and differences in the price level or cost of living. The list of criteria used to compile the index and the weights assigned need to be carefully assessed and thoroughly discussed with all stakeholders to ensure that the main causes for substantial differences in the costs of public service delivery across jurisdictions are captured in the index.

Ultimately, the selection of factors and the assignment of weights are political choices. However, the formulas used to distribute equalization grants should be based on a relatively limited number of factors. The inclusion of too many variables reduces the transparency of the allocation scheme. It is also more costly and difficult to update a larger number of variables on a regular basis, and the use of many variables introduces more opportunities for political manipulation. Therefore a balance has to be struck between simplicity and transparency, and the necessity to find factors that equitably reflect the true fiscal need of local governments. Variables used as factors should accurately reflect needs, come from an independent source, and be free of manipulation by either central government or subnational governments. For example, the use of “inputs” such as school buildings and hospital beds as allocation factors causes inefficiency by providing an incentive that could distort the preferred allocation of resources.

Once the methodology for measuring needs is decided, several steps must be followed to arrive at a practical measurement. We illustrate a five-step procedure to determine a basic index of expenditure needs for each region in appendix table 6.1.1.

Measuring Fiscal Capacity

The next step in the computation of equalization transfers is to determine the level of fiscal capacity, that is, the ability of each region to collect own-source revenues. As in the case of expenditure needs, several approaches can be used to estimate fiscal capacity, which in the case of subnational governments may be defined as the potential revenues that can be obtained from the tax bases assigned to the subnational government if an average level of effort (by national standards) is applied to those tax bases. This means that ideally, tax capacity should be measured by the size of the tax bases available to local governments or the revenue that these tax bases would yield under standard tax rates. Using the actual amount of revenue collections in a locality as a measure of fiscal capacity should be avoided if local authorities can control tax rates, tax bases, or the administrative enforcement effort.

A variety of methods are used around the world to measure local fiscal capacity. One measure is the jurisdiction’s lagged level of revenue collections, for example the revenues collected last year. Using past collections does not satisfactorily address the problem of negative incentives. Sooner or later subnational governments will “learn” that higher collections translate into lower transfers.

Appendix Table 6.1.1 *Computing a Basic Index of Expenditure Needs*

- Step 1. Determine the *aggregate level of subnational expenditure needs* (SEN)—This may be based on historical data on subnational government expenditures.
- Step 2. Select *expenditure needs factors*—These may include population, land area, number of school-aged children, and so on.
- Step 3. Compute each *region's relative need* for each factor (x_i^j / X^j)—For instance, if a region has 10 school-aged children out of a national total of 200, the relative need of this region for this factor is $10/200=0.05$; x_i^j is the size of factor j in jurisdiction i , X^j is the size of factor j in all jurisdictions, and $\sum_i x_i^j = X^j$.
- Step 4. Determine the relative importance of each *needs factor* (a^j)—For instance, if education is 25 percent of the aggregate local budget, the factor “school-aged children” may be assigned a weight of 0.25 in the expenditure needs formula; a^j is the weight assigned to factor j , where $\sum_j a^j = 1$.
- Step 5. Determine the *expenditure need for jurisdiction i* as:

$$\text{Need}_i = \sum_j (x_i^j / X^j) * a^j * \text{SEN}$$

Source: Authors.

Another widely used measure is the per capita level of personal income. This tends to be a good proxy, as it is simple and usually widely available. Yet another widely used proxy measure is gross municipal product (GMP), which is the local equivalent of the national-level gross domestic product. GMP is a more comprehensive measure of the fiscal capacity than per capita income because it includes income generated within a jurisdiction irrespective of where a worker or producer resides.

A modified version of GMP is the Total Taxable Resources (TTR) method, which excludes certain items such as central taxes and transfers to residents since they do not provide a potential tax base.

Finally, some countries (for example, Canada, United States, and Australia) have used a multidimensional measure of fiscal capacity known as the Representative Revenue System (RRS). The basic idea underlying the RRS is to calculate the amount of revenue that a region would collect if it were to exert average fiscal effort. This is done by collecting data on revenue collections and tax bases for each of the taxes under consideration for every locality. Using information on all tax bases for every jurisdiction as well as the national/regional average fiscal effort for each of the taxes, one can compute the amount of revenues that each jurisdiction would collect under average fiscal effort. This amount is then considered to quantify the fiscal capacity of each jurisdiction. The main benefit of the RRS is that computations are made at a disaggregated level and based on detailed knowledge of (proxies for) the statutory tax bases.

Overall, the RRS is a thorough and complete method that accurately measures the fiscal capacity of a region. It is based on disaggregated data and takes into account variations in effective tax rates among various tax components and nontax revenue sources. As a result, fiscal capacity as measured by the RRS can be considered a more accurate representation of a region's true fiscal capacity, compared to that measured by other methods. However, because of the disaggregated nature of the computations, the measure is data intensive and, therefore, it is not always possible to use it.

The practical steps to be used in the computation of fiscal capacity based on a “representative revenue system” are illustrated in appendix table 6.1.2.

Appendix Table 6.1.2 *Computing a Measure of Fiscal Capacity Based on the Representative Revenue System (RRS)*

Step 1. *Select proxy measures for the tax base (T)*—Select measures of a jurisdiction’s own-source revenues (that is, local revenue collections) and a proxy for the region’s tax base, such as aggregate personal income or gross municipal product.

Step 2. *Define the Average Effective Tax Rate (AETR)*—This can be defined as:

$$\text{AETR} = (\sum_i \text{Own-source revenue}_i) / (\sum_i \text{Tax base}_i)$$

This coefficient reflects the average share of own-source revenue in relation to the tax base across all jurisdictions

Step 3. *Compute fiscal capacity*—Fiscal capacity for jurisdiction *i* equals:

$$\text{Fiscal capacity}_i = \text{AETR} * \text{Tax base}_i$$

This amount reflects the amount of collections that each jurisdiction would have if it exerted an average level of fiscal effort in collecting own-source revenues.

Source: Authors.

Measuring the Horizontal Fiscal Imbalance

Having defined expenditure needs and fiscal capacity for each jurisdiction, we can now compute the fiscal imbalance for each jurisdiction. The methodology can be summarized in three simple steps. First, define the *fiscal imbalance as the difference between expenditure needs and the potentially available resources*.¹⁵ The latter include own fiscal capacity, shared revenues, and other transfers. Jurisdictions that do not have a positive fiscal imbalance—in other words, those where potential available resources exceed expenditure needs—are dropped from the process entirely, since they do not qualify for an equalization grant.

The second step is to define the “*relative fiscal imbalance*” for each jurisdiction that has pre-qualified for an equalization grant. The relative fiscal imbalance is simply the relative size of each jurisdiction’s fiscal imbalance as a share of the aggregate fiscal imbalance (of all jurisdictions with a positive fiscal imbalance). The third step is the assignment of equalization transfers in proportion to each jurisdiction’s relative fiscal imbalance (see appendix table 6.1.3).

Phasing In a New System of Transfers

A final important step in the reform of the system of transfers is deciding whether to introduce reforms “cold turkey” or whether to phase in the reforms over a period of years to smooth out the transition to the new regime. The government will have to weigh these considerations if it decides to go forward with the reforms. A more gradual approach would typically make the changes in regional funding more politically acceptable, since it would give the territorial units’

15. This approach measures fiscal imbalance from an additive perspective, by subtracting the potential fiscal capacity from expenditure needs. An alternative approach could be to measure the fiscal imbalance from a multiplicative perspective, where the transfers would be affected by the inverse of the fiscal capacity index. In both cases fiscal capacity is “discounted” from expenditure needs. Depending on the design of the measurement model, both methods can lead to similar results, the former being in general less sensitive to changing assumptions on parameters and easier to understand in terms of its results. The federal government in Russia uses a multiplicative perspective in its equalization formula with the regions.

Appendix Table 6.1.3 *Assigning Equalization Transfers—Compensating for the Fiscal Imbalance*Step 1. *Define fiscal imbalance*

$$\begin{aligned} \text{Fiscal imbalance} &= \text{Expenditure needs} - \text{potential available resources} \\ &= \text{Expenditure needs} - \text{own fiscal capacity} - \text{shared revenues} - \text{other transfers} \end{aligned}$$

For every jurisdiction that does not have a positive fiscal gap, set:

$$\text{Fiscal Imbalance} = 0$$

Step 2. *Define relative fiscal imbalance*

$$\text{Relative fiscal imbalance}_i = \text{Fiscal imbalance}_i / \sum_j \text{Fiscal imbalance}_j$$

The relative fiscal imbalance is the relative size of each region's fiscal imbalance as a share of the aggregate fiscal imbalances of all jurisdictions.

Step 3. *Assign equalization transfer*

Define the equalization transfer to local government *i* as:

$$\text{Transfer to jurisdiction } i = \text{Relative fiscal imbalance } i * \text{Equalization account}$$

Where:

$$\text{Equalization Account} = \text{Regional Transfer Fund} - \sum_i \text{Conditional Transfers}_i$$

Source: Authors.

administrative structures time to adjust. The changes in transfers—and the losers that reform will generate—are often a stumbling block for the implementation of a new system. Drastic changes in the system of transfers and the overall level of funding can hurt local governments' ability to carry through longer-term plans, may create uncertainty, and may increase political friction, destabilizing the reforms.

Basically there are two approaches to phase in a new transfer system. One approach is to introduce the new mechanism gradually, while maintaining part of the status quo for a period of time. In this case, for a number of years the transfers would be the result of a combination of the original allocation mechanism and the new one.

The other approach is to hold jurisdictions partially harmless by partially compensating the losers for their loss in revenues due to the regime shift. Often these approaches fix the base transfer or entitlement for the local governments in nominal terms. The transition to the new system is achieved through real growth and inflation, both of which will contribute to the shrinkage of the historical transfer. This is typically harder to implement at the beginning, because additional funds for the purpose of holding harmless are often not available. Unless there is significant growth in revenues or additional resources could be identified, a hold-harmless provision would significantly reduce the available pool of funds for equalization (that is, the reform implementation may take longer than initially desirable). However, a phased-in program or the use of a hold-harmless scheme may be a preferable alternative to having no reform whatsoever.

In this regard, a viable alternative is to design from the outset a phased-in program of reforms for the system of transfers, aiming at converging to the new system in an explicitly agreed period of time.

Appendix 6.2. “Robin Hood” or Fraternal Equalization Transfers: The Practice of Negative Transfers at the Local Government Level

The systems of transfers for equalization purposes come in many different forms and shapes. For example, equalization funds may be distributed either by formula or on an ad hoc basis. The pool of funds to be redistributed may come from a higher level of government (the central or regional government, as the case may be) or may be directly contributed by the same governments among which the funds are redistributed. This latter system receives a variety of

names, such as fraternal or “Robin Hood” systems. Because the Kozak proposals for subnational reforms in the systems of transfers recommends, among other others, the introduction of negative transfers at the local level akin to fraternal equalization, this appendix reviews some of the international experiences with negative transfers and how well these can be adapted to Russia. We present the experience of six countries (Denmark, Germany, Latvia, Lithuania, Switzerland, and Ukraine) here, three of which are countries also in economic transition and quite close neighbors of the Russian Federation.

Lessons from International Experience

In this section we discuss first the significant dependence of all transfers, including negative equalization transfers, on the design of primary assignments of revenue sources and expenditure responsibilities among the different levels of government, and on the need to address vertical imbalances. Second, we classify the fraternal schemes into several types by source of the redistribution pool. Third, we discuss the mechanism for choosing the extent of revenue equalization.

In all of the surveyed countries except Denmark and Switzerland, the primary revenue assignment results in an extreme vertical imbalance, as subnational governments receive no discretion over the tax rates and bases that account for the bulk of their revenue. Danish subnational governments can set the rates for their major revenue earners (income and land taxes), but at the same time perform many functions that should fall under the responsibility of the central government (for example, social policy). In order to correct this vertical imbalance, general grants (from the central government’s general revenue) are distributed in proportion to local governments’ tax base. In Switzerland, the confederate Constitution gives cantons vast taxing powers enabling them to discharge their functions without being dependent on the federal level. However, subsequent delegation of federal responsibilities to cantons and encroachment by the Confederation into cantonal tax bases have required compensation from the federal government.

The primary vertical imbalance in the four countries other than Denmark and Switzerland is addressed through assigning a share of the central government tax revenue to the subnational level. In Germany, the Constitution prescribes fixed retention rates for the revenue from the personal and corporate income taxes at the point of collection. In addition, a federal law fixes a share of the VAT revenue to be shared with states on per capita bases. In Latvia, the vertical gap is filled with a portion of revenue from the PIT and the entire revenue from the land and property taxes. In Lithuania, it is a portion of PIT only. In Ukraine, shares of revenue from 11 central government taxes are assigned to the subnational level, with PIT accounting for the bulk of the yield. As subnational governments of these four countries cannot on the margin affect their revenue, there is likely to be a mismatch between the revenue yield and the expenditure needs. Unless the revenue sharing is carried out by means of differentiated retention rates, fine-tuning has to be achieved through interjurisdictional redistribution.

Thus, in four out of the six surveyed countries, the bulk of subnational revenue derives from sharing proceeds from central government taxes. Therefore, conceptually, local government contributions to the equalization scheme are funded with the central government’s tax revenue. However, apart from this, in some of the four countries central governments make further contributions to the equalization scheme. In Ukraine, local governments’ contributions are made through daily remittance rates on the assigned tax revenue (dominated by the PIT) while claims from the scheme are covered with daily retention rates on the exclusive central government tax revenue (dominated by the CIT). Therefore, besides the substitution of one flow of cash for another, the central government general revenue makes up the difference between the two. Similarly, in Latvia, the central government makes up the difference between the total contributions to the equalization scheme and total claims on it. Note, however, that in both Latvia and Ukraine, central governments can manipulate the extent of this difference by lowering the base

level of expenditure norm, which would increase contributions to the fund by richer local jurisdictions, lower claims on it by poorer local jurisdictions, and lower the required contribution to the fund by the central government.

In contrast to Latvia and Ukraine, Germany and Switzerland have no central government funding in the equalization scheme, except for the initial assignment of tax-revenue shares to the subnational level. Denmark and Lithuania present a mixed case: their schemes use annually adjusted vertical grants, but these are not related to the balance of contributions and claims in the equalization scheme.

Among the six countries Switzerland is the only one where the size of the equalization pool is fixed (by the subnational share of the three federal taxes). In Germany, Denmark, and Lithuania the pool is determined by the extent of interjurisdictional disparity in fiscal capacity (and, in the Danish case, also the disparity of need). In these three countries, the local governments' contributions have to cover the pool size and are not capped. In Latvia and Ukraine the pool size is determined by totaling the shortfalls of local governments' fiscal capacity below their expenditure needs. However, the local government contributions are capped at 35 percent of the total revenue of the donor jurisdiction in Latvia and at a differentiated limit in Ukraine. Thus in these two countries the difference between the pool size and local governments' contributions is covered by the central government.

Fiscal capacity is directly equalized only in Germany, Denmark, and Lithuania. German federal law prescribes graduated equalization with a target of 92–100 percent of the yardstick (weighted national average). In Denmark, partial equalization of fiscal capacity relative to the average is prescribed in the national law at 80 and 90 percent of the difference from average capacity, respectively, for equalization among counties and among metropolitan municipalities. In Lithuania, fiscal capacity equalization is complete but the target is negotiated annually with the local government association (as of 2000, set at 65 percent of the national average). In Switzerland, local governments' fiscal capacity is not equalized directly but is accounted for with a certain weight in the complex formula of the pool apportionment. In Latvia and Ukraine, fiscal capacity is equalized only relative to the expenditure need of each locality.

In Denmark, Latvia, Lithuania, and Switzerland, the fiscal capacity measure takes into account all sources of local government revenue. In Germany, the consolidated fiscal capacity of a state leaves out 50 percent of revenue from municipal taxes. In Ukraine, the fiscal capacity measure disregards local governments' own taxes altogether and takes into account only assigned tax revenues.

In Germany and Switzerland the horizontal payments are carried out by states, which also administer taxes for all levels of government. In Denmark and Lithuania the equalization scheme is administered by the central governments, which also collect taxes to all levels of government. In Latvia, the payments to and from the equalization fund are administered by the central government's treasury, which also collects taxes for all governments except the cities of Riga and Ventspils. These two cities collect the assigned revenue from the personal income tax themselves and transfer their contribution to the equalization scheme on a weekly basis. In Ukraine, the equalization scheme is administered by the state treasury, which also collects all taxes to all levels of government. However, reportedly the tax administration in Ukraine is under *de facto* dual subordination of both central government and local authorities.

Overall, fraternal grants tend to complement tax-revenue sharing between levels of government. Usually, this tax-revenue sharing aims to rectify vertical fiscal imbalances that cannot be addressed with proper assignment of taxing powers owing to legal, political, or administrative constraints. Tax-revenue sharing does not possess all of the good qualities of real taxing power, however, such as intertemporal predictability of revenue, a stable cash flow, a tax price effect for local services, a revenue link to policy outcomes, and others. Annually budgeted cash grants do not possess any of these good qualities. Different forms of tax-revenue sharing can bring about a combination of some of these qualities—except the tax price effect. Even if local residents face

the tax price of the local services funded with complete retention of the local yield from the shared tax, efficient allocation does not occur because residents cannot respond to this tax price by choosing more or less of the local service.

Tax-revenue sharing can bring about predictability of local revenue if both the total share of assigned tax revenue and the mechanism for its appropriation are fixed and transparent. Local governments have stable revenue flow and accountability for the economic outcomes of their policies if they hold a residual claim on the local yield from the shared tax. However, appropriation of the shared tax revenue exclusively on a derivation basis is not optimal either. Because localities cannot adjust the rate of the shared tax, there is likely to be a mismatch between the distribution of the local yield from the local tax and the distribution of need for local services. If this mismatch were only due to cyclical fluctuations, localities could smooth it out with borrowing (given the predictability of the shared tax yield). In transitional countries, however, underdeveloped markets for municipal borrowing do not allow for this option. On the other hand, differentiation of tax retention rates would hamper the intertemporal predictability of local government revenue. Furthermore, the process of arriving at the differentiated retention rates may be fraught with bargaining, negotiation, and lack of transparency.¹⁶

One significant problem that arises with tax sharing on a derivation basis with uniform retention rates is that it can lead to substantial fiscal disparities when revenue bases and expenditure needs are very unevenly distributed across local governments. One possibility, therefore, is to focus the design of the shared tax system on minimizing the resulting fiscal disparities. This would be accomplished by reducing the uniform sharing rates to the point where only the richest jurisdictions have approximately enough funds to cover their expenditure responsibilities. All the other jurisdictions in this case are financed by equalization grants funded vertically by the central authorities. The needed revenues for the center come from the minimal sharing rates given to the local governments. A completely different approach is to provide local governments with tax-sharing rates that are high enough to fund the expenditure needs of at least a majority of local governments. Of course, in this case the higher homogeneous sharing rates lead to local governments having large potential “surpluses,” while at the same time some other local governments with low tax bases do not have enough resources to cover their expenditure needs. It is in this situation where the implementation of a negative transfers system comes handy. By combining tax-revenue sharing with fraternal redistribution, fine-tuning can be achieved while preserving the predictability and stability of local government revenue flow.

Moreover, fraternal redistribution seems to be especially appropriate for countries in transition, or at least it is not a foreign and distant concept. After all, the traditional regulating approach in the old Soviet budgeting system had many features of a fraternal system of equalization. Local governments with low tax bases and supposedly higher expenditure needs received higher tax-sharing rates and in some cases additional cash transfers. The mechanism used in this system was transparent enough for every one to know that the funds for equalization came from local governments that received much lower revenue-sharing rates and no cash transfers.

Relevance to Russia's Regions

The survey of the experiences in six countries suggests that the fiscal environment in which fraternal transfers operate resembles in many ways the fiscal environment currently observed in Russian regions. Indeed, in Russia, the initial assignment of taxing powers results in an extreme

16. Those were the alleged common features of the “regulated” sharing rates system, which Russia operated during the early years of the transition at the federal-regional level. One way around these problems is to determine the differentiated retention rates by formula. As mentioned above, Vietnam has recently developed this type of approach.

vertical imbalance in the regional-local sector. Moreover, regional authorities have no opportunity to provide local governments with adequate fiscal powers because tax policy is practically monopolized by the federal government. Federal legislation rigidly assigns tax instruments by level of government and does not allow regional governments to devolve part of their taxing powers to localities.

Therefore, in Russia the only two ways for a regional government to pass on fiscal resources to localities are cash transfers and the entitlement to a portion of the yield from regional government taxes. Cash transfers are the least desirable source of local government revenue because of intertemporal unpredictability, uneven flow within the fiscal year, and weak local fiscal accountability. Sharing of the regional tax yield at the point of collection (derivation-basis) provides for more predictability and links a local government's revenue more strongly to the outcomes of its policies.

However, the uniform rate of tax-revenue retention is likely to suffer from a mismatch between the distribution of the shared tax yield and the distribution of expenditure needs. As local governments have no power to adjust the rates of shared taxes, they have no opportunity to balance this mismatch. If the mismatch is addressed with vertical cash grants, that would require a greater centralization of resources, thus hampering revenue predictability and fiscal incentives. If they were annually adjusted, tax-revenue retention rates would be less transparent and lead to negotiation, potential unfairness, and intertemporal unpredictability. However, uniform retention rates combined with fraternal redistribution of the local yield gives enough flexibility to achieve a desirable pattern of revenue allocation while avoiding some of the negatives mentioned above.

In fact, derivation-based allocation is only one of many patterns that can be achieved with formula-based sharing of tax revenue. However, fraternal grants achieve most of these allocations through political consensus. Thus, fraternal redistribution is simply formula-based tax-revenue sharing institutionalized in the form of horizontal consensus seeking.¹⁷ For example, the complete equalization of local per capita tax capacity in Lithuania results in the same outcome as if all assigned revenue were centralized in a pool and then apportioned on a per capita basis.

The international experience suggests that fraternal redistribution is easier (politics-wise) to establish among jurisdictions with similar initial conditions. In this case, interjurisdictional equalization works as an insurance scheme against cyclical fluctuations or structural changes in the economy. Thus, in Germany, the mechanism worked relatively well at the beginning. But even before the unification, political tensions arose in West Germany because the burden of financing the equalization eventually had shifted onto just the two richest states. At unification, the Western states simply refused to accept the new members into the "fraternity," thus leaving the task of equalization of the Eastern states to the federal government. Once a clear distinction emerge between contributors and recipients, it becomes politically difficult to reverse the flow of transfers even when circumstances change.

In cases where jurisdictions are quite heterogeneous, the international experience offers several options to make fraternal redistribution politically feasible. One way is to have separate equalization schemes for homogeneous clusters of jurisdictions (for example, Eastern states versus Western states in Germany, counties versus metropolitan municipalities in Denmark). In the Russian context, this option can be useful for regions that have several types of municipalities (for instance, rayon-level and settlement-level governments, or rural rayons and urban cities). Another way to deal with heterogeneity is to use more elaborated measures of expenditure needs to better take into account the existing differences (for example, higher expenditure

17. Of course, it is not always true that fraternal systems of equalization are consensus based. After all, it is the central government that designs and passes laws on the particular mechanism used.

norms for republican cities in Latvia and larger cities in Germany). Acceptability of fraternal equalization also depends on political alternatives. Even with heterogeneous jurisdictions, richer localities might prefer the fraternal scheme with higher sharing rates over cash transfers from the centralized pool with lower sharing rates. But given the potential political friction that open fraternal systems of equalization can generate, it must be clear that this approach is not expected to be a solution for all Russian regions. For example, political friction could be increased when differences in economic bases are accompanied by a radically different ethnic composition of the local jurisdictions or where local governments differ for other reasons.

The international experience suggests that the size of the redistribution pool should be determined endogenously by the extent of interjurisdictional disparity. However, it is important that the contribution of the central government to the equalization scheme be fixed in advance. In this case, the fraternal scheme works as a consensus builder as the chosen level of equalization would balance the interests of the contributors and the recipients. In contrast, when the regional government fills the gap between contributions and claims, local governments can collude to press for more funds from the regional governments (for example, the cases of Ukraine and Latvia). Thus, if the total yield from the regional tax share assigned to the local level is not enough to fill the aggregate vertical gap, the central governments can complement it with a separate form of vertical grants (for example, per capita grants to Eastern states in Germany or grants proportional to the local tax base as in Denmark).

Although in a fraternal system the ultimate responsibility for funding the scheme rests entirely with richer localities, there should be some safeguards to provide them with revenue predictability and fiscal incentives. Accordingly, if the equalization target is not fixed in the law, it should be limited from below so that the contributing localities would be left with at least that percentage of the regional average of per capita capacity. To provide localities with fiscal incentive the equalization should not be complete but partial, as in Denmark, or graduated, as in Germany. Another option is to leave some portion of local revenues out of the equalization scheme (for example, own taxes in Ukraine and 50 percent of municipal taxes in Germany). This would provide local governments with some incentives to use their own tax instruments (even if they have very limited yield). The obvious candidates are taxes whose bases are not shared with the regional government and thus are hard for the latter to assess.

The six countries whose cases are discussed here present examples of fraternal equalization being administered either by the central governments or the localities themselves—usually the level of government that collects the shared tax. However, in the Russian context that would be the federal government. This might have the benefit of credibility, as localities would know that the regional government cannot renege on its promise during the fiscal year. However, the federal treasury/tax inspectorate might be too rigid to accommodate complex arrangements for interjurisdictional settlements. Thus, fraternal redistribution might require a regional government's own treasury system to operate.

As was explained at the beginning of this appendix, fraternal redistribution is usually used in conjunction with tax-revenue sharing that aims to mitigate the inadequate assignment of taxing powers or compensating for the vertical shift of expenditure responsibilities. This would suggest that fraternal grants are less relevant for regions dependent on federal transfer, where the total regional yield of assigned taxes is not enough to cover the vertical gap. It would be inappropriate to assign all of the regional tax yield (including that from taxes unsuitable for decentralization, such as the CIT, natural resources, and the like) to the local level while keeping the federal transfers in the regional budget. In transfer-dependent regions, cash grants appear to be the only available options for balancing local expenditure responsibility. It is important to note that the distribution of these cash grants should be based on objective criteria incorporated in explicit formulas.

7

Subnational Borrowing

Subnational borrowing, generally considered the fourth “pillar” or dimension of intergovernmental fiscal relations, plays an important role in Russia’s evolving fiscal federalist structure.¹ On the one hand, subnational borrowing is a crucial component of a sound system of intergovernmental fiscal relations because it provides local governments with access to credit, which is much needed for the modernization of crumbling public infrastructure in Russia’s regions and localities. Moreover, rehabilitation of old infrastructure can reduce maintenance costs and thus improve the fiscal position of regional and local governments. On the other hand, there is a need for a balanced policy to prevent irresponsible borrowing by subnational authorities. As such, both the freedom to borrow by local governments and the constraints on this freedom can lead to inefficiencies.

As we show in this chapter, the evolution of the Russian legal framework for subnational government borrowing shows a gradual introduction of constraints on regional and local governments’ borrowing powers. One of the questions we seek to answer in this chapter is whether this policy provides the right balance between federally imposed fiscal safeguards and subnational fiscal autonomy.

Normative Considerations and Lessons from International Experience

On normative grounds, the ability of local governments to borrow as part of the system of intergovernmental fiscal relations is extremely important. The final efficiency of decentralized governance rests on the benefit principle, which requires that the costs of public services be covered with taxes paid by those who benefit from these services. However, if current tax receipts were spent toward some capital project that brings fruits for many years ahead, such as a new road or a water treatment plant, the benefit principle would be violated in an intertemporal sense. It is possible that many of the benefits of a particular capital project would be reaped by those who had not yet been born at the time of construction or moved to the locality years later. Unless corrected, this intertemporal market failure would lead to the underprovision of local public capital goods. Debt financing has the potential to rectify this situation by spreading the costs of building a facility over the entire period during which benefits from this investment are received. Unfortunately, this solution is not safe from an error of the opposite type. Since members of the current generation make the decision about the capital project, they might discount the costs borne by future generations and overspend or otherwise behave fiscally irresponsibly.

Apart from intergenerational fairness, there are other efficiency arguments both in favor of and against subnational borrowing. Even if the benefits of a capital project occur within the lifespan of one generation, it might be desirable to spread its costs over time. Similar to households’ choices, government borrowing can be driven by the desire to smooth consumption; that is, constituents might prefer the accumulation and gradual repayment of public debt as opposed to sudden sharp increases in tax rates.²

1. This chapter draws from Timofeev (2004b).

2. For example, Barro (1979) theorizes that owing to the excess burden of taxation (a second-order effect), an optimal path of debt issue should maintain expected constancy in tax rates.

However, for subnational governments to select the optimal level of capital investment, it is further important that subnational governments expect to bear the full costs of debt repayment. If there is some chance that eventually the debt would be paid out by some other entity, such as the federal government, local governments might have less than due consideration for the costs of capital projects. Not only does the expectation of a bailout weaken the responsiveness of the borrower, it also induces a sort of moral hazard situation by making lenders discount the expected losses. This situation not only results in a failure of local governments to act responsibly, but also in a failure of lenders to subject local governments to credit discipline.

International country experiences in controlling subnational government borrowing present one or a combination of the following four approaches: administrative controls, rules-based control, coordination of borrowing by different levels of government, and reliance on market discipline (Ter-Minassian and Craig 1997). Explicit restrictions on subnational borrowing generally prevail in countries where subnational governments have limited tax autonomy and thus have to rely on funding from the central government (von Hagen and Eichengreen 1996). In contrast, in many federal countries, the states or regional governments are free of borrowing restraints. But even in unitary countries, the rule appears to be that more revenue autonomy is associated with less strict constraints on borrowing (Crédit local de France and Crédit communal de Belgique 1997). However, there are notable exceptions. For example in the Netherlands, where local governments' own sources account for less than 30 percent of their total revenue, subnational borrowing is guided by quite permissive rules rather than discretionary authorization by the national government. In contrast, in Denmark, where almost 60 percent of local government revenue derives from own sources, localities were not allowed to borrow until the mid-1990s and in most cases must still seek approval of the national government.

In countries where subnational governments are free from legislative or regulatory borrowing constraints by higher-level governments, they arguably face borrowing constraints imposed by creditors (Lane 1993). Credit markets can provide important incentives to local governments, such as rewarding fiscally responsible governments with access to credit and lower interest rates; these incentives deter governments from accumulating an unsustainable level of debt. Empirical studies in the United States support this hypothesis by finding a positive relationship between the level of indebtedness of U.S. states and localities and their costs of borrowing (see Bayoumi, Goldstein, and Woglom [1995] on U.S. state bonds; Capeci [1994] on New Jersey municipal bonds). However, the effectiveness of the market-based discipline mechanism hinges on several important conditions—*independent markets for credits, availability of accurate information on a local government's outstanding debt and revenue capacity, and the expectation of no bailout from the higher-level government.* The latter condition seems to be violated in many evolving federations. At the start of decentralization, subnational governments lack capacity for tax administration and thus have to rely on transfers from the central government. If the amount of intergovernmental transfers is determined as the gap between a local government's expenditure needs and revenue capacity (as is generally the case in Russia), the system of intergovernmental transfers can facilitate bailouts if a local government's fiscal deficit is considered to reflect unmet fiscal need. Thus creditors might perceive the higher-level government's financing as an implicit loan guarantee and might not subject local governments to credit discipline.

The credibility of a "no-bailout" policy is affected by factors influencing either the central government's payoff from bailing out a locality or the local government's costs of inducing a bailout. Rodden and Eskeland (2003) analyze fiscal and political institutions in 11 countries, including intergovernmental relations, national politics, and market mechanisms. The functionality of markets for capital, land, and other factors of production determine whether fiscal and market mechanisms can hold local governments accountable for their policies, for example through capitalization of the fiscal outcomes in property values and interest rates. Thus, accountability to the financial markets can increase local governments' costs of fiscal recklessness, which would make

inducing a bailout by the central government a less attractive option. National politics determine the credibility of the central government's commitment to the no-bailout policy depending on how well the creditors' interests are represented at the national level and to what extent the national legislature is prone to log-rolling and vote-trading. Finally, division of authority established in the system of intergovernmental fiscal relations determines the relative responsibility of the levels of government for a local fiscal crisis.

For a bailout to occur, political arrangements should be such that at some point the central government cannot resist the pressure to help out a locality while the local officials can still hold on to their office. This can happen if voters hold the central government politically responsible for local fiscal problems. The central government will be held responsible when it has (or is presumed to have) the financial means to resolve the crisis, regardless of whether the local government diverted or wasted all resources by the time of the crisis, or whether the local government was simply never given sufficient tax autonomy to fulfill its expenditure obligations. Alternatively, the central government might act in the interest of other jurisdictions or creditors (including public employees) rather than the residents of the affected jurisdiction. Furthermore, local government officials should have prior beliefs about the central government's sensitivity to local fiscal crises in order for them to pursue a bailout strategy.

Given its ex post sensitivity to local fiscal crises, it is natural for the central government to attempt to introduce preemptive hierarchical oversight and explicit constraints on subnational borrowing. However, this approach tends to reinforce the belief that the central government is ultimately responsible for subnational fiscal policies. Moreover, as Lane (1993) argues, "the kind of government that is likely to be unresponsive to market discipline, one that persistently incurs unsustainable deficits and limits the information available to lenders, is also likely to seek ways of avoiding surveillance and of circumventing any direct legal limits on its deficits." In short, it is not clear whether the central government would be any better in monitoring local fiscal policy than markets.

The Evolution of the Legal Framework for Subnational Borrowing in Russia

The first law on local self-government of the transition years granted local authorities unconstrained rights to engage in credit operations. The *Federal Law on Local Self-Government* of 1991 allowed local councils to borrow for "social or production" purposes, to participate in the charter capital of commercial banks, to provide budgetary loans to other governments and enterprises, to issue municipal debt for the purpose of investing in social and production infrastructure, and to transfer municipal funds to commercial banks for the distribution of targeted credits to approved projects. In 1992 an amendment to the law abolished the right of local government to provide budgetary loans.

Separate legislation governs legislative oversight of subnational borrowing. In 1993, the *Federal Law on the Foundation of Budgetary Rights* introduced greater oversight by the subnational legislature over the executive's borrowing practices. This legislation allowed the executive branch at each government level to draw intergovernmental loans for short-term liquidity purposes while implementing the budget approved by its corresponding legislative body. In addition, the law permitted the executive branch at each government level to issue public debt for investment purposes under the approval of the legislature, while the executive was also authorized to take out short-term bank loans under general conditions issued by the legislature. The legislation further established that higher-level governments are not responsible for the obligations of lower-level governments.

In 1995, the new federal *Law on the General Principles of Organization of Local Self-Government in the Russian Federation* replaced all previous subnational borrowing and debt provisions with a single vague clause: "Local governments have a right to issue municipal debt and lotteries, to attract and give loans, to establish municipal banks and other institutions of finance and credit

in accord with the laws of the Russian Federation.” However, in 1997, the federal *Law on the Financial Foundations of Local Government in the Russian Federation* introduced a requirement that municipal bonds may only be issued for investment purposes. In this latter law, total volume of municipal debt was further restricted to 15 percent of budgetary expenditures and reiterated that local government debt was explicitly voided of any state guarantee. In addition, the 1997 law banned transfer of municipal property to the charter capital of credit institutions. Among other effects this measure limited the access of subnational governments to credit at below-market interest rates (see box 7.1).

Subnational governments are subject to additional requirements specified in the federal *Law on the Securities Market* of 1996 to the extent that they rely on ruble-denominated securities for subnational borrowing. This law requires the legislative branch at all levels of government to set annual limits on the issuance of public debt. However, further regulation of government securities was left to subsequent legislation. In turn, the federal *Law on the Features of the Issuance and Circulation of State and Municipal Securities* of 1998 established limits on annual borrowing for deficit financing at 30 percent of pretransfer revenues for regional governments and 15 percent for municipalities. The law also required the registration of the issuance of subnational securities with the Ministry of Finance, allowing the Ministry to refuse to register the issuer if it has a record of violating federal laws on budgets or securities.

During 1997–98 a series of presidential decrees authorized 12 regions to issue foreign currency-denominated bonds (so-called Eurobonds—see box 7.2). As a prerequisite for Eurobond issuance these decrees introduced borrowing limits, which were later legislated in the aforementioned federal law on government securities and eventually replaced by the *Budget Code* provisions. While the *Budget Code* outlaws foreign borrowing by subnational governments (as discussed immediately below), a transition clause grants those regions with outstanding foreign debt the right to issue new foreign debt up to the amount of the retiring debt (in other words, refinancing is allowed).

The entire body of budget legislation applicable to federal and subnational governments was replaced by the 1998 *Budget Code*, which actually came into force on January 1, 2000. As opposed to earlier practices, the *Budget Code* prescribes that subnational budget deficits may only be financed with domestic (as opposed to international) borrowing. The *Budget Code* further limits the deficit of local budgets to 10 percent of pretransfer revenues and the deficit of regional governments to 15 percent. Overall, the level of subnational governments’ debt is limited by the amount of pretransfer revenues and by the requirement that expenditures on debt

Box 7.1 *Government Leverage over Subnational Credit Institutions*

Market-based fiscal discipline would be ineffective if subnational governments were able to exercise their leverage on lenders in order to attract credits at below-market interest rates. Many studies suggest that backing from local authorities is a necessary precondition for successful business in Russian regions (OECD 2000). Thus, at all levels of government, commercial banks compete for the right to strike a special deal with authorities and become “authorized” banks. In most cases this special status is very profitable for banks as it enables them to hold budget accounts at little or no interest and obtain government guarantees on loans provided to local producers (OECD 1997). On the other hand, at least at the federal level, there is evidence that *governments* might try to capitalize on their close relations with banks. For instance, in early 1996, the federal government forced several large banks to purchase long-term securities at below-market rates (OECD 1997). Similar practices, even if not documented, might be taking place at the regional level. There is some evidence, however, that those regional governments that sell their securities at local financial markets (for example, the Republic of Tatarstan, Novosibirsk and Chelyabinsk Oblasts) usually face higher costs of borrowing than regions selling bonds at the national markets in Moscow and St. Petersburg (Vedi 1998). In fact, the depressed nature of local financial markets prompts local governments to switch from protecting their own banks to embracing large Moscow-based banks, which are actively penetrating regions.

Box 7.2 Eurobonds

Presidential Decree No. 304 of April 8, 1997 instructed the federal Ministry of Finance to authorize the issuance of foreign currency–dominated bonds by the City of Moscow, City of St. Petersburg, and Nizhny Novgorod Oblast. The decree required the Ministry of Finance to develop federal standards for issuing such securities. The authorization of Eurobonds was conditioned on meeting the 30 percent limit on the ratio of a region’s annual borrowing to its pretransfer revenues and the 15 percent limit on the ratio of a region’s annual debt servicing expenditures to its pretransfer revenues. The decree also required the borrowing regions to obtain credit ratings from at least two international agencies.

The City of St. Petersburg was the first Russian region to sell these bonds on international financial markets. According to Vedi (1998) the May 6, 1997 issue of five-year Eurobonds, totaling US\$300 million, was sold at 99.5 percent of the face value with an annual coupon of 9.5 percent (at a spread of 312.5 basis points to the U.S. Treasury bond yield). One-third of the proceeds from St. Petersburg’s first issue of Eurobonds was deposited to commercial banks for up to six months. On May 27, 1997, the City of Moscow issued three-year Eurobonds totaling US\$500 million. This issue was sold at 99.8 percent of the face value with an annual coupon of 9.5 percent (at a spread of 315 basis points to the U.S. Treasury bond yield).

Another Presidential Decree (No. 1,128, October 18, 1997) extended the conditions of the first decree to cover the Moscow Oblast, Sverdlovsk Oblast, and Krasnoyarsk Krai. The third decree (No. 1,212, November 12, 1997) added the Republic of Komi, Leningrad Oblast, Oryol Oblast, and Samara Oblast to the list of regions allowed to issue Eurobonds. The third decree also approved the establishment of an “independent federal rating agency” (*Interfax Rating Agency*) and instructed the Government of Russia to use the agency’s ratings when evaluating the regions’ applications for authorization of securities issuance. The fourth and fifth decrees (No. 1,297, December 17, 1997, and No. 280, March 20, 1998) authorized the Chelyabinsk Oblast and Republic of Tatarstan to issue Eurobonds on the same conditions.

The last decree in this series (No. 696, June 9, 1998) introduced additional conditions for the registration of Eurobond issues. Regions were required to include data on their budget revenues, expenditures, and deficits along with financing sources for the three preceding years in their application for authorization of Eurobond issues. In addition, the last decree established that federal authorization of subnational borrowing would be guided by the priorities of federal borrowing needs. As noted earlier, these conditions were later legislated in the federal *Law on the Features of the Issuance and Circulation of State and Municipal Securities* and eventually replaced by provisions in the *Budget Code*.

service do not exceed 15 percent of total expenditures in any fiscal year. These limitations on subnational fiscal deficits and debt in Russia are somewhat similar in size and scope to the restrictions imposed by the Maastricht Treaty for European Monetary Union (EMU) members. For instance, given the share of subnational governments in Russia’s total public sector, the limits on annual subnational deficits are comparable in size to the Maastricht rules, which allow member states to run a deficit of 3 percent of GDP. However, limiting the total stock of outstanding debt of Russia’s subnational governments by the size of annual flow of their pretransfer revenue (10.5 percent of GDP as of 2001) appears much more restrictive than the Maastricht rules, which allow public debt up to 60 percent of GDP for EMU member states.

Subnational government borrowing is restricted to financing capital expenditures; that is, recurrent deficits are not allowed in the *Budget Code*.³ Subnational government debt is unconditionally secured with all the assets of the borrower. The *Budget Code* requires all levels of government to maintain a debt ledger. The debt ledger of a higher-level government contains information on the debt of all of its subunits. The *Budget Code* specifies that failure to comply with these legislated requirements and borrowing limits could result in a budget audit or in the assumption of control over budget execution by the higher-level government.

3. Enforcing a “strong golden rule”—that government borrowing should only be used to finance capital investments—would require separation of recurrent and capital budgets. In contrast, a “weak golden rule” limits the size of a subnational government’s fiscal deficit by the amount of public investments.

Up to now this provision has been deployed only as a political threat. As the legislation does not cover procedures of government bankruptcy and forced administration, higher-level authorities do not know how to exercise this power. Some regions (such as Samara Oblast) have attempted to introduce the procedure in their regional legislation, but thus far there has not been a single precedent of using this procedure. In 2001, seven regional governments failed to meet the legislated limits on borrowing without any repercussions from the federal government (*Vedomosti*, December 5, 2001). Recently proposed legislative reforms by the Kozak Commission specify modalities for the forced administration of regional finances by the federal government.

In summary, the current federal legislation limits annual deficits of subnational governments, the volume of outstanding debt, and the form and purpose of borrowing at both regional and local levels. Thus, over the last decade Russia has shifted from a permissive stance on subnational borrowing which solely relied on market-based discipline to administrative and rule-based controls. The administrative controls take the form of discretionary registration (which effectively means approval) of subnational securities issuance with the federal Ministry of Finance. Besides introducing a “weak golden rule” on subnational borrowing, federal legislation attempts to induce quasi-market sustainability of subnational debt by linking the volume of borrowing to the revenue capacity of subnational governments and relating debt-servicing outlays to the size of their budgets.

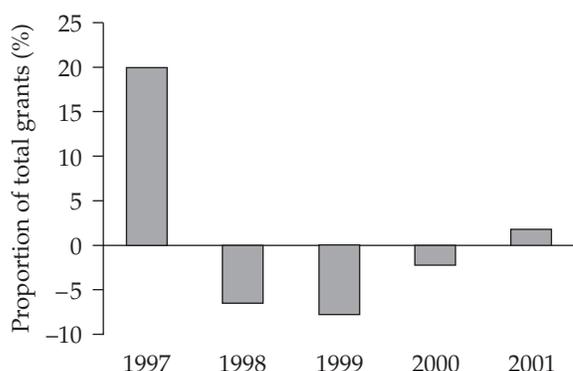
National Politics and Subnational Borrowing

Apart from intergovernmental arrangements that emphasize federal preeminence in subnational finance, the credibility of the officially stated “no bailout” policy in Russia is further hampered by the residual political culture of undivided authority. There is a great difference between the legislated system of government and Russia’s political reality. Although the 1993 Constitution establishes federalist principles, the idea of one-man management is still common for political rhetoric at the national level. For example, in his inauguration speech, President Putin stated: “In Russia, the head of state always has been and always will be a man responsible for everything that happens in the country” (*Time Europe*, May 22, 2000). Furthermore, subnational leaders often pay lip service to this rhetoric, especially when it comes to accepting responsibility for local woes. Such ambitions make the central government vulnerable to the strategic moves of local authorities. On the one hand, the Constitution constrains the federal government’s powers to discipline subnational officials. On the other hand, when local government mismanagement makes national news, everyone in the country turns to federal authorities for a response. The unclear division of responsibilities among government levels, in terms of legislation as well as voter perception, makes the central government partially responsible for local fiscal outcomes.

While normally the federal government tries to respect the fiscal autonomy of subnational authorities, it cannot help but *intervene* in urgent situations as a result of political considerations. Political pressure for the federal government to step in becomes particularly strong when subnational governments accumulate huge arrears in wages to public employees or payables to utility suppliers. For instance, during the 1996 presidential campaign, the federal government provided loans to regional governments to clear wage arrears in subnational budget entities (OECD 2000). Complete clearance of subnational wage arrears occurred only in 1997, when the volume of federal loans exceeded 20 percent of the explicit federal grants

Figure 7.1 shows the flow of federal loans as a proportion of federal grants in the aftermath of the 1997 bailout. Although the budget loans are partially repaid in the years immediately following, the underlying perverse incentive to seek additional budget loans is clearly visible in the upward trend. Not surprisingly, by 1998 the volume of wage arrears accumulated by subnational governments had approached the 1996 level again, causing the federal government to roll its loans over to the regions while introducing certain conditionalities. Among these was

Figure 7.1. The Flow of Federal Loans as a Proportion of Federal Grants



Source: Authors.

one-to-one matching of regional funds allocated for clearing arrears and dedicating a fixed share of regional budgets to wage payments.

Such federal interventions arguably create new perverse cash management practices at the subnational level by which regional and local governments prefer to allocate funds to trivial needs because urgent needs can be funded with a federal bailout later. One example is the accumulation of wage arrears to subnational employees (such as teachers and local healthcare workers) on the eve of national elections. Other opportunities to squeeze the federal budget arise in relation to one-time seasonal expenditures. For example, deliveries to certain areas of the far north regions can be made only during the summer. If the supply of necessities is not replenished before those regions become inaccessible, the population of those areas faces a winter of hardship. Similarly, maintenance of district heating systems has to be completed by autumn, otherwise pipes would be damaged by frost and their repair would require much larger outlays. Such deadlines allow subnational governments to build political pressure and play a budgetary game of “chicken” with the federal government.

A telling example of how subnational governments exploit opportunities to squeeze funds from the federal government is provided by the annual heating crises that occur in the *Primorskii Krai*. Every summer local authorities in the *Primorskii Krai* appear to have no cash reserved to stock up on fuel and carry out essential repairs to their heating systems (*The Washington Post*, February 6, 2001). Given that subzero temperatures are a certainty for that part of the country, the central government eventually provides extraordinary funds to the region in order not to risk the safety of millions of local residents. This game culminated in a major crisis during the winter of 2000/2001, the coldest winter since 1949. At the peak of the crisis about 100,000 residents in a dozen *Primorskii Krai* localities were living in unheated homes as decrepit boilers and pipes cracked and burst throughout the region. In addition to making extraordinary cash transfers, the central government had to airlift pipes and plumbers for emergency repairs to heating systems. Besides the direct expenses, this heating crisis caused significant political costs to the central authorities. National media channels routinely showed residents of the region huddling around electric heaters as ice formed on the walls of their apartments. In the end, the Kremlin used its political muscle to “persuade” the region’s governor to take responsibility for the crisis and step down, thus limiting the political damage inflicted on the President’s reputation by a crisis that left him looking weak and ineffective.

The example of the *Primorskii Krai* illustrates the contradiction between the great ambitions of the central government and the limited vertical controls provided by the Constitution. This contradiction is endemic to the power struggle among government levels unfolding since the collapse of the Soviet Union. While the 1993 Constitution reflects the balance of power prevalent during the early transition (during which the regions appeared to have the upper hand),

this balance has evolved substantially over time in favor of the federal authorities. Despite the political autonomy granted to subnational governments in the Constitution, the central government has effectively used the recentralization of taxing authority as well as reforms of the vertical political power structure to gain the political upper hand in intergovernmental relations. However, the price of this federal dominance has been the absence of hard budget constraints at the subnational level.

In summary, the political culture of undivided authority, reinforced by the rhetoric of national politicians, prompts the public to hold the central government responsible for every function, regardless of what level of government is formally responsible for carrying it out. Moreover, the centralization of taxing authority at the federal level weakens the credibility of the central government's no-bailout commitment because the central government is perceived to be responsible for any local fiscal woes—and it possesses the means for rescuing the troubled jurisdictions.

Estimates of Outstanding Subnational Debt

Before the *Budget Code* required subnational governments to maintain debt ledgers, the extent and variation in subnational debt had been hardly documented. Yet, given the permissive legislation on subnational borrowing before the introduction of the *Budget Code*, many authors had long suspected that there was a significant growth of subnational debt in Russia (Lavrov, Litwack, and Sutherland 2000; Martinez-Vazquez and Boex 2001; OECD 2000). This section examines the data on outstanding debt reported by subnational governments to the Ministry of Finance since 2000 in accordance with the new legislation. Before proceeding to analyzing the recent data, we survey previously available evidence on subnational borrowing in order to validate the data and identify recent trends.

In the absence of comprehensive data on subnational borrowing and debt, Freinkman, Treisman, and Titov (1999) assessed the development of subnational debt in Russia based on data from various sources. According to their estimates, the stock of subnational commercial debt (that is, debt other than intergovernmental loans) grew from 0.11 percent of GDP in 1994 to 1.51 percent in 1997. One-third of this amount accounted for loans from commercial banks and slightly more than one-quarter was estimated to arise from subnational securities. The local share of consolidated subnational debt was estimated to be about 20 percent. According to Freinkman's estimates, as of 1997, the total subnational debt in the Russian Federation was estimated to be about 6 percent of GDP if unreported liabilities were taken into consideration. For example, overdue payables to employees and suppliers (in other words, budget arrears) grew from 1 percent of GDP in 1995 to 2.58 percent in 1997. Although the estimated level of indebtedness was not high by international standards, Freinkman, Treisman, and Titov expressed concern at the time about the rapid growth, high yields, and short maturities of the subnational debt instruments used in Russia.

Similar figures for 1998 appear in OECD (2000), which was based on partial survey information from 53 out of a total 89 Russian regions. This study estimates the stock of subnational commercial debt (foreign and domestic) to stand at about 4.46 percent of the gross regional product (GRP). Intergovernmental loans accounted for an additional 1.49 percent of GRP. Finally, subnational government guarantees on loans to third parties accounted for 2.76 percent of GRP. In the OECD study, about 46 percent of subnational government guarantees for domestic commercial loans and securities are reported to be in arrears. Overall, the stock of subnational governments' debt and guarantees, both domestic and foreign, was estimated to be at 8.4 percent of GRP. Two-thirds of this amount is attributable to regional governments while localities make up the rest. Regional governments' borrowing makes up 90 percent of foreign subnational debt (1.55 percent of GRP). Foreign borrowings by local governments are dominated by on-lending programs from the World Bank and the European Bank for Reconstruction and Development (EBRD).

In 2000, the first official figures on the volume of outstanding debt appeared in the subnational government reports prepared in accordance with the norms introduced by the new *Budget Code*.

Although the reports do not provide a breakdown of intergovernmental loans, commercial loans, and securities, subnational debt can be broken out by regional and local government levels. As reflected in table 7.1, the total stock of subnational governments' (domestic and foreign) debt, guarantees, and overdue payables was reported to be 2.44 percent of GDP in 2000 and 2.48 percent in 2001. These figures are rather modest and roughly consistent with the trends reported earlier by Freinkman, Treisman, and Titov (1999) and OECD (2000), given the three-fold ruble devaluation of 1998.

However, there is a danger that these and earlier estimates might be subject to underreporting by subnational authorities, as these authorities have significant incentives to disguise the actual stock of their explicit borrowing and guarantees given the federal ceiling imposed on the level of the latter portion of debt.⁴ For example, in 2000, only 55 out of 88 regions (excluding the Chechen Republic) report non-zero values for total indebtedness (explicit debt and guarantees on loans to other parties). We are more inclined to trust the figures reported for the overdue payables, as those are used by subnational authorities as a bargaining tool to squeeze aid from the federal government and thus are less likely to be underreported.

As a data validation exercise, we consider the reported subnational debt stock in table 7.1 with the sum of the annual flows of deficit financing reported by subnational governments in the past. Although debt ledgers were not introduced until 2000, regions have been reporting data on subnational deficit financing (credit loan minus redemption) since 1995, although separate figures for regional and local levels are not available before 1997. Hence, table 7.2 presents

Table 7.1 *Reported Stock of Outstanding Subnational Debt, 2000–01, as a Percentage of GDP*

	<i>Local</i>		<i>Regional</i>		<i>Total Subnational</i>	
	2000	2001	2000	2001	2000	2001
Foreign debt	0.01	0.01	0.40	0.46	0.41	0.47
Domestic debt	0.11	0.16	0.40	0.75	0.50	0.91
Total explicit debt	0.12	0.17	0.80	1.21	0.91	1.39
Domestic guaranties	0.07	0.08	0.36	0.40	0.43	0.48
Overdue payables	0.87	0.48	0.23	0.14	1.10	0.62
Total debt (Explicit, contingent, and implicit liabilities)	1.05	0.73	1.39	1.75	2.44	2.48

Source: Calculated from Russian Ministry of Finance data.

Table 7.2 *Accumulated Flow of Domestic Financing of Subnational Deficit since 1994, as a Percentage of GDP*

<i>Financing method</i>	1995	1996	1997	1998	1999	2000	2001	2002
Intergovernmental loans	0.07	0.31	1.05	0.81	0.52	0.50	0.53	0.56
Securities	0.07	0.22	0.43	0.37	0.17	0.08	0.14	0.28
Other	0.25	0.64	1.13	1.81	1.17	0.93	0.63	0.54
Total	0.39	1.17	2.61	2.99	1.86	1.51	1.30	1.38

Source: Calculated from Russian Ministry of Finance data.

4. For example, an audit by the Federal Chamber of Accounts revealed that the Government of Magadan Oblast concealed 63 percent of the actual stock of debt outstanding at the end of 2001. In fact, the actual stock of debt exceeded the federal ceiling (the amount of pretransfer revenues) by a factor of eight. Although the results of the audit have been sent to the federal executive and legislative authorities, it is not clear whether any action will be taken in the absence of legally established procedures.

the accumulated flow of subnational borrowing by the consolidated regional-local sector since 1995.⁵ Consistent with our hypothesis that local governments may be underreporting their debt portfolio, the calculated amount of subnational domestic debt using this methodology (estimated at 1.51 percent of GDP) is three times the stock of domestic debt reported by subnational governments in 2000 (0.50 percent of GDP). The flow of deficit financing suggests that at its peak in 1998, the accumulated domestic debt could be as high as 3.00 percent of GDP. Apart from intentional concealment, underreporting of debt stock could also be explained by the lack of proper records before the introduction of debt ledgers. Indeed, according to table 7.1 the reported stock of explicit domestic debt increased from 0.50 percent of GDP in 2000 to 0.91 percent in 2001. At the same time, the flow of deficit financing suggests that during 2001 the stock of domestic subnational debt rose by only 0.03 percent of GDP and actually decreased relative to GDP because of economic growth and inflation. Therefore the rest of the rise in the reported debt level could potentially be explained by better accounting for debt accumulated in the past.

The budget classification before 1999 did not record foreign borrowing as a form of deficit financing. However, from the deficit financing figures we know that the net flow of foreign borrowing to subnational governments was +0.004 percent of GDP in 1999, -0.6 percent in 2000, and -0.5 percent in 2001. This means that in 1998, the stock of foreign borrowing by subnational governments could have been as high as 1.1 percent of GDP. Overall, the total amount of explicit subnational debt, domestic and foreign, could have been more than 4.1 percent of GDP in 1998. This is only slightly more than one-third of the annual pretransfer revenue of regional and local governments.

By all estimates, Russia's subnational debt falls below the levels typically observed in other countries. This is especially noticeable in comparison with the four major federations: total subnational debt equals 57 percent of subnational expenditures (or 11 percent of GDP) in Australia; 91 percent (21 percent of GDP) in Germany; 100 percent (25 percent of GDP) in Switzerland; and 65 percent (13 percent of GDP) in the United States (Martinez-Vazquez and Boex 2001). Out of the EU countries with comparable size of the subnational sector, only four countries have smaller subnational debt relative to their GDP: Ireland (1.5 percent of GDP), Denmark (3 percent), Italy (3 percent), and Finland (3.5 percent) (Falzon, Peretti-Stahl, and Verdier 2002). At the same time, in Sweden, the United Kingdom, Spain, and the Netherlands, subnational debt exceeds 5 percent of GDP. It should be noted, however, that in all EU countries except Luxembourg the ratio of local governments' share of total public debt to their share of total public spending varies from 1/7 in Italy to 2/3 in France, whereas in Russia this ratio is 1/7 (if social security is not included in the equation).

Despite this modest average indebtedness, however, based on 2002 figures, a dozen of Russia's regional governments accumulated a stock of debt exceeding their annual pretransfer revenue. Moreover, in the Republic of Tyva, the stock of debt exceeds pretransfer revenue by a factor of four. As mentioned before, it is not clear how the federal government is going to enforce the borrowing limits in the absence of procedures to enforce the federally legislated requirements. Moreover, in the case of Tyva, where half of the accumulated debt is attributable to federal loans, the problem of lack of legal enforcement tools seems to be compounded by a lack of political will to enforce a hard budget constraint.

Data on the composition of deficit financing offers some insights into the dynamics of subnational borrowing by debt instrument during 1995–2002. In all years except 1997, loans from banks and enterprises, reported as “other borrowing,” constituted the major source of credits for subnational government.⁶ Intergovernmental loans were the second-largest category of subnational

5. Note that our calculations may suffer from double counting of federal loans that are in turn lent by regional governments to the constituent localities.

6. In this respect, Russia is closer to the EU countries, where taking bank loans is more common for local governments than issuing bonds (as is the case in the United States or Canada) (Swianiewicz 2003). However, recently bonds have been issued more intensively by European local governments as well.

borrowing throughout these years and the major one in 1997 (when the federal government undertook a campaign of clearing wage arrears). The share of securities in the annual subnational borrowing was below 20 percent throughout the period 1996–2001. However, securities accounted for 42 percent of subnational borrowing in 1995 and made a comeback again in 2002, accounting for 56 percent of subnational borrowing in that year.

The Role of Subnational Securities

The estimates of the outstanding stock of subnational debt presented above suggest that securities played a relatively minor role in subnational borrowing for most of the transition period. While the first issues of ruble-dominated subnational bonds date to 1992, the real volume of issuances did not grow significantly until 1994 (Lykova, Krylova, and Chebanova 1999). The largest volume of subnational securities, amounting to 0.85 percent of GDP, was issued in 1997. Over one-third of this amount was accounted for by the City of St. Petersburg's issuance of bonds. Another quarter was accounted for by "agrobonds," which regional governments used as payment to the federal government for defaulted agricultural credits that the regional government guaranteed (see box 7.3).

In 1997, the yield on high quality subnational securities was 5–10 percentage points above that of federal bonds (RET 1997). Starting in 1998, the volume of subnational securities steadily decreased. In June 1998, on the eve of the financial crisis, rising sovereign risks led to a three-fold increase in the yield on subnational bonds and to a significant decline in credit supply. After the financial crisis, the annual issuance of bonds had been consistently below the volume of retired securities until 2001. As can be seen from table 7.2, in 1998, subnational governments switched from issuing securities, first to attracting commercial loans and later to borrowing from the federal government. However, in 2002 security issuance once again became the preferred form of subnational borrowing.

For completeness' sake, we should also mention the role of promissory notes and bills of exchange (*veksels*) issued by subnational authorities, which flourished briefly in the mid-1990s (for a good review, see OECD 1997). In addition to being a debt instrument, these also constituted a convenient means of exchange for complicated barter schemes. As a result of massive nonpayments at that time, many legal entities had their bank accounts attached. Yet, unlike cash, *veksels* provided the security of moving through the stages of the agreed scheme without being withheld at the request of the federal tax authority or other creditors. However by 1998, promissory notes and bills of exchange constituted less than 1 percent of subnational domestic debt (OECD 2000). Such a small volume should be attributed to the enactment of the 1997 *Law on Bills of Exchange and Promissory Notes*, which banned government bodies from issuing *veksels* (OECD 1997).

Proposed Reform Measures

The recommendations of the Kozak Commission concerning subnational borrowing and debt aim to address one of the most critical issues in the federal oversight of subnational borrowing, notably the absence of an effective mechanism for the federal government to enforce its own legislation.

The Kozak proposals would effectively introduce legal procedures for enforcing hierarchical (administrative and legislated, rule-based) constraints on subnational governments' liabilities. The proposed legislative reforms would allow for the temporary financial administration of a jurisdiction's finances by the higher-level authority for a period of up to one year.⁷ The forced administration of local government finances by a region can be introduced by the regional arbitration court

7. Besides financial administration, this provision allows the higher-level governments to take over powers of the constituent jurisdiction in other circumstances, such as natural disasters, calamities, and misuse of delegated state powers.

Box 7.3 Agrobonds

Agrobond securities were issued by 69 regional governments during 1997–98 as a means to cover regional guarantees on commodity credits provided by the federal government to local agricultural producers in 1996. As Russian farmers viewed government credits as another form of subsidy, not many of them intended to repay the credits and thus regional government guarantees became due. Some regions opted for paying to the federal government upfront the credits that they had guaranteed, but the majority converted them into securities hoping that the federal government would not press too hard for payment. However, the federal government auctioned off these bonds to private investors, at a discount of about 20 percent. Out of almost RUR7 billion in agrobonds, roughly one-third ended up held by foreigners, with another one-third held by Russia's largest banks and investment firms and the rest (unsold) held by the federal government.

Each region made three agrobond issues of equal size: one-year, two-year, and three-year bonds. All bonds had the same nominal value (RUR10,000) and an annual (or two semiannual) coupon of 10 percent. It follows from this that all issues of agrobonds should have been redeemed by June 2001. However, by December 2001 only 42 regions had paid off the principal and the interest on the one-year issue. Moreover, only 27 regions paid off the principal and both coupons of the two-year issue. As for the three-year issue, only 24 regions paid the principal and all three coupons. Twelve regions did not make a single payment on any of the agrobond issues.

In July 1998, agrobond holders complained to federal authorities about the regional governments' delay in payments. In response, the federal Ministry of Finance declined to intervene, citing the budget autonomy of regions. The only concession made by the federal government on the issue was easing the regions' debt burden by restructuring regional liabilities on the portion of agrobonds that remained in federal ownership.

With the exception of the Karachayevo-Circassian Republic, none of the 12 regions whose agrobonds failed to find buyers at the auctions and thus remained completely in federal ownership have made a single payment on these liabilities. The Karachayevo-Circassian Republic presents an exception as the liabilities of its agrobonds were offset against special federal budget appropriations for drought aid to the Republic in 1998.

Out of 26 regions whose securities were completely sold out at the auctions, only 18 had cleared them all by December 2001. Many regions whose agrobonds were partially sold in 1997 have cleared the securities held by private investors while refusing to service the portion held by the federal government. This can be explained in part by the fact that some law firms buy defaulted agrobonds for pennies on the secondary market and aggressively seek settlement through courts with varying degrees of success. In addition, regional governments' debtors often buy agrobonds at a significant discount to pay off their debt to regions at face value. It should be mentioned that the ruble devaluation of 1998 significantly reduced the real value of agrobonds.

Source: Ministry of Finance (Letter No. 05-01-04 of May 6, 1998) and the Rosbank Custody Department.

with respect to a defaulting *municipality* and by the Highest Arbitration Court with respect to a defaulting *region*. The necessary legal ground for such a ruling would be constituted by actions (or lack of actions) on the part of the subnational jurisdiction leading to overdue liabilities in excess of 30 percent of its own-source revenue, provided that higher-level governments do not have overdue liability to this jurisdiction. The Kozak proposals also introduce a legal responsibility of subnational governments for their state (municipal) institutions.

In order to restore the financial capacity of a defaulted jurisdiction, the temporary financial administration is supposed to take measures to restructure the overdue debt. The temporary financial administration can prepare changes and additions to the current year budget as well as for the next fiscal year. These changes and additions to the budget have to be approved by the legislative body of the defaulted jurisdiction. In some cases defined by federal law, the changes and additions to the current and tentative budgets have to be approved by the higher-level legislature as well. The temporary financial administration also ensures proper execution of the budget and takes over other functions of the local government if necessary.

Conclusions

All available estimates—and actual budget reports since 2000—indicate that the accumulated stock of subnational debt in the Russian Federation is not high relative to international standards. One possible way to explain this is to interpret it as evidence of the effectiveness of market-based discipline before 2000. Indeed, when the volume of outstanding debt approaches the level that cannot be served in unfavorable states of the economy, lenders charge the borrower a higher risk premium and thus discourage further borrowing (Lane 1993). At some level of indebtedness the borrower will not be offered further credit at any interest rate and thus will be denied access to credit markets. As such, the adequate functioning of credit markets could have contributed to limiting local government debt. However, this explanation seems to be contradicted by the contention that credit markets in Russia were weak during the early transition and that the necessary data to inform subnational credit decisions were generally unavailable prior to 2000.

An alternative explanation for the modest levels of subnational government indebtedness is the tightness of the credit market in Russia. Until the financial crisis of 1998, tight monetary policy combined with excessive borrowings by the federal government drove the real interest rate up to double-digit values. If the low average level of debt were in fact caused by the tight monetary policy (that is, the high cost of credit), it would indicate that at least the demand-side component of the market-based fiscal discipline worked and local governments indeed responded to high interest rates. Another piece of evidence in support of this hypothesis is the significant share of subnational total debt attributed to overdue payables to employees and suppliers. It is documented that private firms responded to the tight credit markets with extended payment periods and, in this way, received (involuntary) trade-credits from their suppliers (Alfandari and Schaffer 1996; Commander, Dolinskaya, and Mumssen 2000). It is possible that subnational governments at least tried to employ a similar strategy.

Although the low average debt of subnational governments might suggest that market-based fiscal discipline might have actually worked in Russia, we are unaware of any rigorous studies that would support this hypothesis. In fact, the observed low subnational debt levels might simply have resulted from the absence of adequately functioning credit markets for subnational governments during the early transition. However, the lack of hard evidence one way or another did not preclude the federal government from changing the budgetary legislation in favor of more legislative and regulatory constraints on subnational borrowing. In fact, the imposed constraints aim to achieve quasi-market outcomes. Similar to market forces within a credit market, the new federal legislation seeks to constrain subnational governments' borrowing within the subnational government's capacity to service debt by relating the maximum debt ceiling to the annual volume of pretransfer revenues. The question is whether federal authorities can enforce these constraints better than the markets. Moreover, discretionary registration of subnational securities issuance with the federal Ministry of Finance might stand in the way of shifting subnational borrowing from attracting bank loans to more transparent bond-financing.

For the market-based discipline to work, several important conditions should be in place: free markets for credits, availability of information on outstanding debt, and no realistic expectation of higher-level bailout. It has been argued in the literature that when those conditions are not present, the federal government cannot be presumed to do any better in monitoring local fiscal policy than the markets (Lane 1993). Thus, there is a potential that the substitution of the market mechanism with administrative controls could actually result in poorer performance of local borrowing and credit markets in the future. An alternative policy direction could be to try to strengthen the market forces, for example, by enforcing disclosure standards or eliminating vertical fiscal imbalances. However, the jury is still out on this issue, and the new government rules may not weaken but rather strengthen market forces in the longer term. As a compromise, in the medium term, higher-level governments could facilitate local government borrowing through municipal lending agencies until the private market for municipal debt is more fully developed.

8

Budgeting Practices and Fiscal Management at the Subnational Level

While policy reform during Russia's transition has encompassed virtually every sector of the economy, modernization of *public sector budgeting practices* and *fiscal management techniques* has been widely recognized as being the most important to the success of the economic transition. The blatant inefficiencies in public production and fiscal imbalances that became evident by the end of the Soviet era clearly implied that the top-down planning system could not be relied upon for the efficient delivery of public services. Therefore, the transformation of the public sector necessitated reforming the budget system to one that is more compatible with the very different role of the public sector in a market economy.

However, the breadth and pace of fiscal management and budgeting reform in the Russian Federation, as in other transitional countries, largely fell short of expectations during the first decade of the transition. The urgency in dealing with fiscal challenges led to a reliance on short-run solutions. Thus, shortfalls in revenue collections were initially met with adjustments in the money supply and with increased government borrowing later, without fundamental policy changes and cuts in expenditure commitments. Eventually, Russia's poor budget practices and "living beyond its means" led to the propagation of budget arrears at all levels of government and payment arrears to the rest of the economy. Moreover, the social pressure to exempt certain items (such as wages) from the budget sequestering implied that budget cuts had to be more severe on other items (such as materials and maintenance), further delaying and also multiplying budget requirements for future years. Rather than reprioritizing expenditure programs, scarce fiscal resources were spread too thin for funding existing (and even a few new) programs, which resulted in the low productivity of public expenditures.

In this chapter we review fiscal management and budgeting practices at the subnational level in the Russian Federation. We also take a look at best practices and principles in fiscal management around the world to identify policies and reforms that could be introduced to improve budgeting at the subnational level in Russia. As part of our discussion on budgetary practices in this chapter, appendix 8.1 considers the use and accounting of extra-budgetary funds in greater detail, while appendix 8.2 provides a brief overview of global trends in public budgeting.

The Budget Process in Russia's Subnational Governments

The budget processes at the regional and local levels of government in Russia are prescribed by the regional (local) laws of each respective jurisdiction. However, regional and local laws on the budget process have to comply with federal requirements introduced in the *Budget Code* of 2000. The *Budget Code* assigns to the executive branch at each level of government the exclusive responsibility for budget formulation and execution and assigns to the legislative (representative) branch the responsibility of considering, approving, and controlling of the budget. While the *Budget Code* improved the uniformity of budget practices at the subnational level and ensures a certain minimum standard in budget practices, budget approaches and expenditure management techniques at the subnational level often leave much to be desired and significantly lag behind federal reforms.

Several remnants of the previous regime's budget practices play important roles in the budget process of regional and local governments. In many subnational jurisdictions, the preparation of the budget is interrelated with the preparation of the socioeconomic development plan, which is often drawn up by the subnational planning department—typically a successor of the local arm of the former State Planning Ministry (*Gosplan*). In addition to providing macroeconomic forecasts, this plan generally proposes a list of “targeted” programs aiming at achieving desired policy outcomes. While many see this as yet another relic of the central planning era, the *Budget Code* continues the approach of relying on a set of minimum social standards as the basis for budgeting. Officially, these financial norms—which should be issued by the federal government and possibly refined down the administrative hierarchy—are binding for all lower-level governments. However, as of 2004, these minimum norms are yet to be developed.

A noticeable institutional feature of subnational budgeting in Russia is the relative status of line departments with respect to the department in charge of budget formulation (typically the Finance Committee). This is yet another remnant of the Soviet era. In the old system, the central plan determined the resource allocation and line departments were in charge of budget preparation and implementation in their respective sectors (Diamond 2002). The Finance Committee played the passive role of consolidating the sectoral budgets and accounting for the allocations carried out by line departments. In the execution phase, the Finance Committee acted as a central funding agent disbursing funds to line departments so that they could ensure production of the planned outputs. This practice has been abandoned at the federal level, where the Ministry of Finance has gained preponderance over the Ministry of Economy (the successor of the central planning agency) and the line ministries. However, in modern Russia, many regions and localities follow a mix of procedures for the budget process, in part those inherited from the common Soviet past, in part established according to the recent *Budget Code*, and in part copied from federal practices.

Despite variations in budget practices across regional and local governments, the budget calendar followed by most subnational governments is fairly routine (Kuzmin and others 2002). Below we summarize the main steps of the subnational budget process including budget formulation/approval and budget execution/control along with their approximate timing in the annual subnational budget cycle.

Budget Formulation and Approval

Budget formulation and approval involves the following steps:

- Around July, preceding the budget's fiscal year, the regional government provides local governments with initial figures for drafting local budgets (including minimum social and financial standards, projected sharing rates of regulating revenues and revenue estimates used in calculating those rates, and regional adjustment coefficients to minimum social and financial standards).
- Around August, local line departments provide the local Finance Committee (or its equivalent) with estimated costs of public service provision and implementation of targeted programs in the coming year. Revenue forecasting practices of local Finance Committees vary from adjusting past collections for expected inflation to surveying production plans of major businesses in the locality. Based on these estimates and the initial figures received from the regional government, the local Finance Committee submits a preliminary draft budget to the regional Finance Department or the comparable regional executive body in charge of budget formulation. In the event of unbalanced revenues and minimum expenditure needs, local governments can also submit requests

to the regional governments for changes in sharing rates of regulating revenues, as well as subsidies and subventions.

- The regional Finance Department (or its equivalent) reviews the draft local budgets, and makes proposals for changes and adjustments. In the event of outright or partial rejection of those proposals by local governments, the regional authorities draw up a protocol of disagreement. They may also set up a Conciliatory Commission for discussing disputed issues, taking into account the targets of the oblast and local budgets.
- In September the regional executive body—based on the review of local budgets by the regional Finance Department—establishes targets for its consolidated regional budget and provides information on the targets of the consolidated and oblast budgets to the federal Ministry of Finance and to the regional Duma.
- Around October, the regional governor presents his budget message to the regional Duma. This includes the (regional and local governments') budget forecast, regional spending priorities, the proposed basic principles of fiscal relations between the oblast and local governments, and information on the fiscal sufficiency of local governments.
- Around November, the local Finance Committee or its equivalent adjusts the preliminary draft budget in line with the revenue sharing agreed with the regional government.
- Around December, the local executive head presents his budget message to the local Duma, after which the local Duma considers and approves (or rejects) the draft budget. The draft budget is prepared following the national budget classification. The general parameters include the total volume of planned revenues, expenditures, and deficit. Revenues are forecast for each tax assigned by federal legislation to the local level as well as for some local non-tax sources. Expenditures are broken down into current and capital categories and classified by categories and subcategories of the functional classification, targeted programs, and by "chief funding agents"—local line departments and other large local spending agencies. The draft budget also specifies the sources for deficit financing and limits on the outstanding debt for the coming year.

Budget Execution and Control

Budget execution and control involves the following steps:

- Within 10 days of the passing of the local budget, each "chief funding agent" within the local government determines quarterly appropriations of funds for the lower-level "funding agents" and also for final spending units in the immediate subordination of this "chief funding agent."
- Within 17 days of the approval of the local budget, the local Finance Department consolidates information on budget appropriations by "chief funding agents," approves it, and sends it to the department in charge of budget execution (either the local treasury or the Finance Committee itself) as well as to the representative body and auditors.
- Within 10 days of the approval of the budget appropriations, the local treasury (or the Finance Department, when also in charge of budget execution) sends its notifications down the hierarchy of "funding agents" to the final spending units.
- Within 10 days of receiving the notifications, spending units prepare their own budgets and submit them to the higher-level "funding agent."
- Within five days of receiving the budget of a spending unit, the respective "funding agent" approves it and sends it to the department in charge of budget execution.
- Based on the approved appropriations and the state of revenue collections, the local treasury (or the Finance Department) issues limits on expenditure commitments by "funding agents" and spending units for a period of up to three months.

- Within the authorized limits, spending units make payment orders to the local treasury, which after checking conformity with the appropriations and budget limits, authorizes the payment from the single budget account. The decision on the authorization of payment has to be made by the local treasury within three days of receiving the payment order. In localities without a treasury system, “funding agents” and spending units have their own bank accounts, which are credited by the higher-level “funding agent” throughout the fiscal year. In this case, the department in charge of budget execution cannot control transactions of the spending units and has to rely on ex post reporting.
- According to the *Budget Code*, the local treasury (or the local Finance Department) does not have discretion to set limits on expenditure commitments different from the approved budget appropriations, except in cases of deferring and blocking expenditures for a particular “funding agent” or spending unit. The deferral cannot affect more than 10 percent of total government expenditures and exceed the term of three months. The local treasury can block expenditures for a particular “funding agent” and spending unit if the initial budget appropriations had a certain conditionality attached and this was not met by the recipient by the time of issuing limits on expenditure commitments or submitting the payment orders.
- “Funding agents” have the authority to alter approved budget appropriations for the subordinated “funding agents” and spending units provided that for each recipient the alterations do not exceed 5 percent of the initial appropriations.
- The initial budget appropriations can be changed by the regional or local executive body in case of revenue shortfalls provided that the amount of changes does not exceed 10 percent of the total budget. Moreover, the alterations cannot exceed 10 percent of total appropriations for any “funding agent” and 15 percent for any spending unit. More extensive changes to the initial budget appropriations have to be backed up with amendments to the initial budget law approved by the legislative body.
- On December 31 of each year the single budget account is closed and the balance, along with revenue and expenditure accounts, is reported to the representative body in the final report on budget execution. Before consideration in the representative body, the final report has to be externally audited.

Two main themes are noticeable in the budget process for regional and local budgets since the introduction of the *Budget Code*. First, the reformed budget process still leaves local budgets highly dependent on the regional budget process. This reduces the effective level of fiscal autonomy at the local level and is in sharp contrast to the situation faced by regional governments. Since the establishment of the grant formula in 2000, federal transfers have become more predictable and can be estimated in advance by regions.¹ Moreover, the federal budget estimates become known to regional governments after the budget draft is submitted to the federal parliament in September of the preceding year. The second main theme in the subnational budget process is a developing treasury function at the local-regional level that is considerably strengthened against past practices. This should contribute to more uniform fiscal discipline at the local level and therefore more stability and lower deficits at the regional level. Before the passage of the *Budget Code* and the beginning of the introduction of modern treasury systems at the subnational level, budget execution had failed to control expenditures at the commitment stage, while controls at the payments stage only led to the accumulation of significant budget arrears during the 1990s.

1. The grant formula undergoes some adjustments on an annual basis but those are vented at the Standing Task Force on IGFR including representatives of about 30 regions and disseminated among other regions around May of the preceding year. In addition to these refinements, the only federal input to the regional budget planning is the forecast of inflation and GDP growth.

Budget Accounting and Reporting at the Subnational Level in Russia

The most controversial issues remaining in Russian budgetary reform at the subnational level are those pertaining to budget accounting and reporting. This section discusses some of the main problems with the existing subnational budget process.

The new *Budget Code* requires that the bookkeeping by all levels of government exclusively follow the uniform chart of accounts sanctioned by federal authorities. In addition, the Russian budget system continues the tradition of bottom-up reporting on budget execution. Every quarter governments at all levels report to the higher-level government budget figures consolidated from those reported by lower-level governments. Since 1995 all levels of government have been using the same centrally introduced budget classification based on the United Nations' System of National Accounts (SNA). Expenditures are reported by function, by economic character, and by "chief funding agents." For every "chief funding agent," expenditures are also classified by object. The revenue side of the budget reports receipts for each tax allowed by federal legislation and also gives some information on non-tax revenues. Until 2000 budget reports did not provide information on the stock of outstanding debt. However, the deficit-financing section reported the flow and composition of annual borrowing. Starting in fiscal year 2000, the budget reports have been providing information on the outstanding stock of domestic debt, foreign debt, and issued guarantees.

Some subnational governments have protested the obligation, in a federal country such as Russia, to report budgets and executions to the higher-level governments. The experience around the world in federal countries is quite varied, ranging from complete freedom in the use of budget classification and no reporting requirements to much more stringent requirements on the other extreme. The provisions in the *Budget Code* and current practices put federal-regional arrangements for budget accounting and reporting in Russia pretty much at the middle of the range. A uniform budget classification system for all levels of government has very significant positive externalities for basic fiscal information flows, policymaking, and good governance. For example, systematic and uniform budget reporting facilitates the comparisons of performance across subnational governments. Quarterly reports of budget execution may be a bit demanding in comparison to similar requirements in other federal countries—one solution would be to change this requirement from quarterly to annually. However, in truth this requirement does not impose extra burdens on regional governments, which in any case have to prepare similar reports for their regional legislatures. What is more important is that the federal authorities have no control or approval rights over the regional budgets or their execution. But as we already have seen, the situation is quite different for the reporting requirements between local and regional governments. In this latter case, regional authorities still exercise control and effective approval rights over the budgets of local governments. This practice of Russia's constitutes an extreme that it is not observed even in some unitary countries.²

The *Budget Code* adopts "publicity" as one of the fundamental principles of Russia's budget system. This principle requires the promulgation of budget plans and execution reports, the completeness of the information provided, and openness of budget procedures to citizens and the media. The principle stands in stark contrast to the secretive nature of government budgets and budget processes during the Communist era, a practice that continued during the transition. Unfortunately, however, at the subnational level the publicity requirement has been often reduced to the publication of a summary of revenues and expenditures in a local newspaper. But since 2000, the federal Ministry of Finance has been publishing quarterly reports on the execution of subnational budgets on its Web site. These reports provide figures on regional governments' revenues, expenditures, and liabilities, including explicit borrowing, payables in

2. Interestingly, in 2001 Russia's neighbor Ukraine, a unitary country, abolished all regional government rights to supervise and approve local budgets.

arrears, and issued guarantees. In addition, the Ministry Web site reports all registered issuances of subnational governments' securities.

The accuracy of budget reports at the subnational level historically has been affected by the presence of budget arrears and extra-budgetary funds. Despite the fact that the new treasury function, now being introduced in many subnational governments, controls expenditures at the commitment stage, budget execution at the subnational level is still reported on a cash basis in Russia, thus excluding receivables and payables accumulated by governments. However, in the late 1990s many subnational governments started reporting accumulated payables to the higher-level government in a separate file according to the same classification used for actual outlays. With the adoption of the new *Budget Code* in 2002 all subnational governments have been required to maintain a ledger of purchases registering terms of order and information about the supplier.

The main problem with budget reporting had been one also of concealment of full information and disclosure. Private lenders and the federal government alike faced the same issues of possible deliberate concealment of pertinent budgetary information by subnational governments through extra-budgetary transactions. Similar to practices observed in other countries, Russian subnational governments for a long time used extra-budgetary funds as a way to avoid fiscal discipline but also as a possible way to protect their funds from claw-back by the higher-level government. The general characteristic of extra-budgetary funds is that some portion of public funds is operated in off-balance-sheet activities (Lane 1993).

There are currently two major channels for off-balance-sheet transactions: public-enterprise funds and extra-budgetary funds (OECD 2000). Although budget accounts and the accounts of public entities are controlled by the executive and, thus, are closely interlinked, budget reports cover only the first component.³ For instance, receipts from user charges for public services are not included in budget reports as they are earmarked for and collected by public entities delivering these services. Also, borrowing by public enterprises is not included in the government budget although it constitutes an implicit obligation of the government. In particular, guarantees for commercial loans to enterprises constitute a considerable share of subnational governments' obligations that had been absolutely unaccounted for before the introduction of the *Budget Code*.

There are two types of extra-budgetary funds: territorial branches of the national funds (also called "earmarked funds") and locally established funds. The latter type is formed with budget allocations unspent in the past year, "voluntary" contributions from local enterprises, tax penalties and other fines, and proceeds from para-governmental ventures. There is no reliable information on the extent of these locally established funds; however, partial surveys indicate a minimum of 2 percent of total budget revenues (OECD 2000). By contrast, resources of the earmarked funds amount to 15 percent of the budget. As of 2001, two-thirds of this amount was covered by the revenues of the road fund, financed by the road user fee.⁴ In 2000, as prescribed in federal legislation, subnational governments consolidated all earmarked funds into their general government budget.

Although subnational government finances are subject to federal standards of budget accounting, they are not safe against manipulations. Subnational officials still have the ability to channel public moneys through extra-budgetary funds and public-enterprise accounts. Thus, some government activity is inevitably excluded from operations covered by the budget. Besides hampering vertical accountability, these practices also interfere with budget comprehensiveness and thus preclude the allocation of fiscal resources based on overall spending priorities. This might be too high a price to pay for allowing extra-budgetary accounts to exist.

3. This is common practice in many countries. See, for example, OECD (1997).

4. This fee used to be 1 percent of enterprise turnover. In 2003 the road user fee was replaced with the transport tax levied according to engine power.

There is little doubt that if claw-back by the higher-level government makes local officials hide funds in extra-budgetary accounts, it would serve the national interest to abandon this type of intergovernmental behavior. In addition, to the extent that extra-budgetary funds serve as informal ways for the local governments to raise additional revenue, their regularization should coincide with giving localities the power to legislate what is now collected in the form of “voluntarily contributions” from businesses as formal levies. This would greatly improve the transparency of local finances. In short, there are significant advantages to eliminating all extra-budgetary accounts at the subnational level, the most important of which is putting an end to the current fragmented approach to budgeting and expenditure-policy formulation (see appendix 8.1).⁵

There is not a great deal of information on how subnational governments actually have adopted the new legislation on budget reporting. One Russian research center carried out the evaluation of how subnational budget practices conform to the IMF fiscal transparency standards (Zamiatina and others 2002).⁶ Although the survey covered only three cities and five regions, the results nevertheless provide some insights into the current state of fiscal transparency at the subnational level. In most localities, the general score was determined as “fair” but that resulted from higher points on the legal framework and lower points on the accessibility of the budget process to non-government stakeholders. The respondents acknowledge some accessibility of budget information for the general public but at the same time the report notes the lack of opportunities for public input to the budget process. In fact, public hearings on policy issues are still very rare at the subnational level and the first ones have taken place only recently.

The situation might change in the near future if the Kozak Commission’s proposals are actually implemented. In particular, article 52 of the proposed law on local government requires public hearings on both the budget draft and the budget execution report. Moreover, the proposed law requires publicizing the budget draft, the final report on budget execution, and quarterly reports on budget execution, public employment, and payroll expenditures.

Other Shortcomings of the Budget Process at the Subnational Level

In addition to extra-budgetary funds, several other items in the subnational budget process require attention. In this section we review the most salient issues.

Program-Based Budgeting

Most subnational governments use a very simple line-item budgeting system inherited from the Soviet era, when local budgets were centrally planned. In many subnational governments, the Finance Department still exhibits a lack of proper status, which arises from its former subordinate function as the financing arm of the central planning office. This reduced status is reflected in the importance of the “sectoral budget branches” within the budget structure, which deal with the budgetary financing of major economic sectors. During Soviet times, budget preparation by line departments and spending units were based on the rigid use of physical norms (linked to line items rather than activities and programs). In the absence of central norms, line departments currently merely roll over expenditures of the previous years, which creates a general absence of economic analysis of alternative resource allocation.

5. It is interesting to note that other transitional countries, such as Kazakhstan and Ukraine, have eliminated all extra-budgetary accounts at the subnational level.

6. The St. Petersburg Center of Humanities and Political Studies’ “Strategy” surveyed representatives of government, media, and NGOs in the cities of Petrazavodsk, Pskov, Velikije Luki, and five subjects of the federation (the City of St. Petersburg and Murmansk, Samara, Novosibirsk, and Yuzno-Sakhalinsk regions) in 2001.

International trends in expenditure management have been moving away from line-item budgeting and toward program-based budgeting (appendix 8.2). There have been attempts to address emerging needs through the adoption of “targeted programs,” which are budgeted on top of line items. These programs are often implemented by the same departments whose traditional expenditures are budgeted in the line-item section. In general, there is a potential for increasing efficiency by explicitly formulating the traditional activities of line departments in the same way as for the “targeted programs.” This would allow local governments to revise priorities in terms of programs and allocate scarce funds according to these priorities. However, currently, there is little indication of any established trends to use performance budgeting at the subnational level in Russia.

Subnational Treasury Functions

Those subnational governments that have not adopted a modern treasury function still suffer from inefficient cash management and the lack of proper control over expenditures. After the funds are disbursed to line departments, the finance office of these subnational governments loses all control until it receives the ex post budget execution reports. In particular, the budget offices in these subnational governments have no control over the expenditure commitments of the final spending units. This creates problems in ensuring that the overall spending limits are not breached. Timely—or “flash”—reports from line departments tend to be overly aggregative, and only in cash and not on a commitment basis. Quarterly reports go into great detail but the information comes with substantial time lags. Moreover, multiple accounts for both revenue and expenditures in the banking system slow down cash transmittal through numerous tiers of intermediary accounts until it finally reaches spending units. This also leads to uneconomic accumulation of cash balances.

Lack of Dialogue

There is a lack of dialogue between the executive and legislative branches over budget policy especially at the budget-formulation stage. This is a legacy of the time when local Soviets were supposed only to rubberstamp centrally planned budgets. Even now detailed reports submitted to the legislature often do not go beyond the minimum standards of budget accounting required in federal legislation. In some localities, local finance offices report more information on local budgets to the higher-level government than to local councilors.

External Audit

External audit is rather weak in Russia, as historically there were no external auditors in local Soviets. In jurisdictions where the representative body recently has established an agency in charge of external audit (typically called the Chamber of Accounts) the latter lacks capacity—as well as clarity on its role in the various stages of the budget process. In the executive branch, local Finance Committees inherited control subdivisions from the Soviet era (called KRUs). However, these subdivisions operated as investigative rather than preventive institutions. They carried out special investigations on alleged irregularities and fraud rather than internal audit to aid the department’s budget management. Internal audit activities focus mainly on compliance with laws and regulations; there is little consideration of value-for-money auditing.

A particularly controversial provision of the *Budget Code*, and one that is reaffirmed in the proposals of the Kozak Commission, is that higher-level governments are allowed to audit the accounts of “transfer dependent” jurisdictions. The center’s position is that since “transfer dependent” jurisdictions use funds provided by federal or regional governments, those higher-level governments should be able to scrutinize how those funds have been used. The position

of “transfer dependent” jurisdictions, on the other hand, is that they are discriminated against for the simple reason that they are poor. It would seem that the positions on both sides are somewhat justified. Perhaps the answer lies in how this issue is commonly handled in other countries. Higher-level governments are justified in auditing the use of conditional grants or transfer funds that are earmarked by the higher-level government to a particular program or use by the lower-level government. Unconditional grants, such as equalization transfers, are meant to be general-purpose funds for local governments, the same as general revenue sharing, and are not supposed to be audited by the higher-level governments. If this latter practice is adopted, poorer local governments would be put on a level basis with richer jurisdictions, which may not get equalization grants but are likely to get both conditional grants and revenues from tax sharing.

Overall, the presence of decentralized institutions, local elections, some decentralization of resources, and established rules of financial accounting and budgeting have all failed so far to lead to accountable budget policy at the local level. Apart from local councils there have been no other mechanisms of interaction between local governments and their constituents that would allow the latter to express their preferences and hold decisionmakers accountable. Citizen participation in the budget process is a particular challenge because of the centralist traditions of public finance and public administration in Russia. There is no tradition of public participation in fiscal decisionmaking (public hearings, consultations, or referenda), or appeals concerning government actions when grievances are justified. During the Soviet era, civil society was suppressed and the Communist Party programmed the few existing nongovernmental organizations (NGOs). This created a culture of mistrust of any public activity. As a result of all this, people in Russia are still extremely unwilling to participate in any joint effort to influence local governance. Overcoming this history no doubt will take some time but there are positive steps that can be taken by legislating requirements for open local councils, public hearings, consultations, referenda, and citizen appeal processes.

Options for Further Policy Reforms

The Soviet era left Russia’s regional and local governments with budget practices that were hardly efficient in their days and totally incompatible with the new role played by subnational governments in a market-based environment. Over the last decade some progress has been made in reforming Russia’s budget system including the subnational budget processes. This progress is especially noticeable in the legal framework introduced with the *Budget Code* in 2000. However, implementation of the new system is likely to take time and currently there are still a number of problems with budgeting and fiscal management at the subnational level in Russia. Unlike many other issues in intergovernmental fiscal relations, budget management at the subnational level was hardly addressed by the Kozak Commission and no significant reforms in this area are considered at the moment.

Nevertheless, our brief review and analysis above lead to several recommendations for further reforms of fiscal management at the subnational level:

- The budget paradigm at the subnational government level needs to shift from control to performance. This process needs to be slow and deliberate and accompanied with granting greater discretion to governmental units for shifting funds among items only after adequate mechanisms are introduced to hold budget managers accountable.
- The budget or finance departments at the local level must be given full authority to prepare the local budget under the guidance of the executive government and approval by the local legislature.
- The adoption of a modern treasury function by all local governments should be a short-run priority for reform.

- There is a need to find vehicles that would increase dialogue between the executive and legislative branches over budget policy especially at the drafting and external audit and evaluation stages. External audit and more in particular ex post budget evaluation need to be strengthened for all subnational governments in order to ensure expenditure prioritization and horizontal accountability.
- External ex post audit by higher-level governments should not be conducted exclusively for “transfer dependent” jurisdictions but rather for all lower-level governments that receive earmarked funds or conditional grants. Funds received from shared taxes or from unconditional equalization grants should be treated as general local government funds and not be subject to mandatory external audit by higher-level governments.
- All extra-budgetary accounts at the subnational level should be eliminated, thus enhancing the comprehensiveness of the budget and improving public-resource allocations.
- Federal policymakers should seek avenues to increase interaction between subnational governments and their constituents and greater accountability for local governments by legislating open local councils, public hearings, consultations, referenda, citizen appeal processes, and the like.

Appendix 8.1. The Use and Accounting of Extra-Budgetary Funds

A long-time area of concern in fiscal management in the Russian Federation has been the lack of comprehensiveness of the budget process, and the impact extra-budgetary funds have on the ability of the subnational government to engage in sound fiscal management practices. Of course, this concern has not been unique to the Russian Federation. For instance, in response to similar concerns, Kazakhstan recently moved to consolidate all of its extra-budgetary funds into its main budget account.

There is a broad consensus among fiscal experts, both in the Russian Federation and abroad, that the existence of extra-budgetary funds is an undesirable fiscal practice.⁷ Thus the question becomes not *whether* all special accounts and extra-budgetary funds should be integrated into the regular budget process, but rather *how this can be achieved* without excessively stifling the flexibility that is necessary for managing certain government operations. Having a comprehensive budget approach does not mean that all expenditures should be managed according to the same set of procedures. For efficiency, specific arrangements for administering some programs may be established, provided that they do not lead to a fragmented approach to budgeting and expenditure-policy formulation.

This appendix provides an overview of how (extra)budgetary funds are used and accounted for in the budget processes of a number of large, developed market-based economies. In particular, we reviewed, for comparative purposes, the (extra)budgetary practices in three countries—the United States, Australia, and the United Kingdom. The review of international practices seeks to highlight a number of important issues. For instance, the money flow: do resources flow in and out of the government’s main account through the treasury or do agencies maintain their own resources and bank accounts? When a government unit engages in

7. Indeed, the crucial role of extra-budgetary funds in achieving a sound budget system is highlighted by the IMF’s Code of Good Practices on Fiscal Transparency. Even though the Code only consists of five pages, it contains seven references to extra-budgetary funds. These references concern the inclusion of extra-budgetary funds in the definition of the public sector (paragraph 1.1.1); executive accountability (1.1.4); compliance with all budgetary regulations and administrative rules (1.2.1); comprehensive coverage of the national budget document (2.1.1); budget classification (3.2.1); reporting of extra-budgetary out-turns (3.4.1); and audit and parliamentary control over extra-budgetary funds (3.4.2).

business-like activity, does it have control over its own receipts? How is the agency's level of control over its receipts determined? Furthermore, what is the chain of executive accountability, what special budgetary controls and audit rules are in place, and what are the reporting requirements for (extra)budgetary funds?

Although each of these countries has a unique budgetary approach, there are many commonalities in their accounting practices. We find that different countries have found different solutions to the question of extra-budgetary funds. Two of the countries reviewed (Australia and the United Kingdom) opted for full inclusion of all (extra)budgetary activities in the main budget account. Other countries (for instance, the United States and Canada) have chosen to place some funds "off-budget." Based on the experience of the countries reviewed, we present lessons for budgetary reforms in the Russian Federation.

Defining the Issues: Budget Coverage and (Extra)Budgetary Funds

A first task in the discussion of extra-budgetary funds is to define the scope of the issue and establish a common vocabulary. What do we mean by the term, "the government"? What is "the public sector"? What are "(extra)budgetary funds?" And how do the two latter concepts relate to the scope of the government's budget?

The term "general government" is typically considered to refer to the combination of all government units, which includes central government as well as state and local governments (Schiavo-Campo and Tommasi 1999). In addition, depending on their size and the way they are managed, sometimes social security funds are classified as a separate subsector of the general government. Alternatively, social security funds are sometimes simply included in the central government budget.

In addition to the government's operations, the public sector includes entities owned or controlled by the government, such as state-owned enterprises or financial institutions. In market economies, public enterprises should be commercially oriented and aim or work as "for profit" entities. For example, while a municipality may own and operate a local public utility (such as a municipal power company), in order to be efficient and profitable, this public utility should have a high degree of autonomy, which is best achieved by incorporation. Therefore the expenditures and revenues of publicly owned companies should not be submitted to the same scrutiny and approval mechanisms as is the government budget, which should cover only the financial transactions of the individual government unit (for instance, the central government).⁸

Comprehensiveness of the budget process raises special issues for government agencies that, while truly part of the general government sector, engage in activities that have business-like features. Specifically, in the normal course of business, some government units (such as public hospitals or national parks) may collect user fees or other revenues. How should such receipts be managed in the context of the overall budget process? In many developed countries a significant share of government expenditures is managed through special budgetary procedures while keeping these accounts within the overall financial management framework of the government. Special budgetary arrangements for such special budgetary funds or accounts include special funds for specific expenditure purposes (such as road funds or health funds) managed at the sectoral level; emergency funds; budgets of autonomous or decentralized agencies, which are especially common in the higher-education and health sectors; special accounts managed by the Ministry of Finance or its Treasury Department; and revolving funds, which are used to provide more flexibility for business-like activities of certain government agencies in order to overcome the need for annual allocations. While in most developed countries expen-

8. Dividends or transfers to/from such enterprises would obviously be included in the government's budget.

ditures managed through such arrangements are typically included in the regular budget, in many developing and transitional countries such accounts or funds are instead managed through extra-budgetary arrangements.

Extra-budgetary activities could be broadly defined as financial transactions (revenues and expenditures) that occur outside the scope of a government's regular budget process and fall outside the government's main budget account. Sometimes extra-budgetary funds are included in the regular budget documents (albeit outside the main government account). In many cases the parliament only has indirect or less-than-discretionary control over extra-budgetary activities. Since budgetary practices vary greatly from country to country, the term "extra-budgetary funds" does not capture a uniform set of government activities.

Establishment of special budgetary arrangements to accommodate diverse government activities might involve a number of dimensions, including special appropriations arrangements (such as standing appropriation or earmarked revenue sources), more independence in control and decisionmaking (for instance, through an independent board as opposed to political appointees), and greater managerial flexibility and reduced government oversight.

There are arguments both in favor of and against including these types of activities and budgetary funds in the regular budget; two philosophically different views on how trust funds relate to the budget frame the current debate. According to *off-budget* proponents, the trust fund mechanism represents a contract with the taxpayers to spend revenues on the specific activities identified as the purpose of the trust fund. For instance, automobile owners who pay an excise on gasoline have a right to expect a share of these receipts to be earmarked for the construction and maintenance of roads. It follows that the special-purpose nature of these funds means they are not discretionary budget resources, and therefore their special character would give reason to exclude them from the regular budget accounts. In the view of those who support the *unified budget* approach, a dollar (or ruble) of federal revenue is a dollar of federal revenue. Spending decisions, therefore, need to be in the context of national rather than programmatic requirements. According to this view, the principle of budget comprehensiveness suggests that such activities should be included in the regular budget process since such funds—like all other revenues and expenditures—constitute government resources and outlays. This view holds that the proper budgetary priorities can only be identified when all potential use of resources compete with one another and the winners get funded from a common budget pool.

The discussion on whether certain funds should be "on-budget" or "off-budget" is not only driven by administrative or economic considerations. The determination to leave or move a fund "off-budget" often has significant political dimensions as well:

- In many countries, the extent of coverage of the government budget determines the size of important fiscal policy measures, including the size of the budget deficit. Modification of the definition of on-budget expenditures might artificially make it easier to achieve legislated or policy-determined fiscal deficit targets. In that case, the government could manipulate overall permissible spending levels by moving accounts either on-budget or off-budget. For instance, if the social security fund is collecting more than it is paying out (owing to demographic dynamics, for instance), inclusion of the social security account on-budget could hide a large structural fiscal deficit. Alternatively, by moving accounts off-budget, one could increase spending for these purposes outside the regular budget process without formally increasing budgetary expenditures.
- Political special interests could seek extra-budgetary status for sectoral funds to secure "protected" funding for sectoral activities, since in many cases off-budget funds are financed through earmarked revenues and have greater discretion over receipts arising from business-like activities.
- Special interests could seek extra-budgetary status for sectoral funds to achieve increased control over expenditures or less scrutiny over special funds.

- Regarding social security funds, questions of equity and solvency are important political considerations. Proponents of keeping these funds off-budget might feel that it would be easier to “rob” the social security fund if the fund is an on-budget account. The overriding issue here is ensuring that those paying into the fund now will receive benefits at some point in the future, whether the fund is on-budget or extra-budgetary.

A Comparison of Approaches and Lessons Learned from Other Countries

The budgetary approaches in United States, Australia, and the United Kingdom with respect to extra-budgetary funds reveal several key differences, but also highlight a number of important similarities. Based on the review of international practices, we draw lessons for the Russian Federation in four key dimensions of the use and accounting of extra-budgetary funds:

- What accounting mechanisms could be used for special accounts and funds?
- How could the appropriation of special or extra-budgetary funds be determined and what level of discretion should fund managers have?
- How could accountability, budgetary control, and audits be dealt with?
- How should extra-budgetary funds be included in the cycle of budget reporting?

ACCOUNTING MECHANISMS The spectrum along which extra-budgetary funds are defined is by the degree that its accounts are integrated into the governments’ system of accounts. In other words, the further a budget fund is outside the regular set of government accounts, the harder it is for the government to control the fund’s resources through its financial management process. All three budget systems reviewed here have essentially integrated all extra-budgetary activities into their government accounts, whether it is in the main budget account (the United Kingdom and Australia) or whether it is outside the main budget account yet still within the system of government accounts (the United States).

Regarding the flow of resources, the dominant accounting approach in the countries surveyed is to take in revenues and receipts through the treasury system and then credit the resources to a special account within the consolidated (general) fund or to a special fund outside the consolidated fund but still within the system of government accounts. Regarding outlays from special funds, there is a greater variety of approaches. In the case of the United States, all outlays of budgetary funds are made through the treasury and then debited to the special account, thus allowing more central control. In the case of Australia, agencies have their own bank accounts and are primarily responsible for reporting on what are called “special accounts.”

It is key for a sound fiscal management system to contain all budget accounts within the system of government accounts, either inside or outside the main budget fund. The main lesson from the international review is that greater flexibility at the agency level is only possible with a highly sophisticated integrated financial management system such as the Australian approach. It follows that it would be preferable in countries with lower administrative capacity (including transitional countries such as the Russian Federation) to contain all budgetary activity within (the special accounts of) the general budget fund. Requiring that all receipts flow through the general treasury account while requiring that all outlays be also made through the treasury would allow former extra-budgetary activities to be managed, audited, and accounted for effectively within the government’s financial system.

APPROPRIATION AND BUDGET DISCRETION All countries reviewed in this study recognize that certain government activities are of a special nature that does not fit in a one-size-fits-all budgetary approach; every country makes a distinction between normal budgetary expenditure items (from the general budget fund), “special accounts” for special purposes such as social security (often with earmarked resources), and “revolving” accounts for business-like govern-

ment operations. However, the manner in which each country determines which status to accord to budget accounts is different.

In the United States, the designation of a special account, budget fund, or revolving accounts is allowed by legislation passed by Congress. In the United Kingdom, the Chancellor of the Exchequer prepares a special request for Parliament in a Treasury Minute, seeking approval for a government department to retain certain revenues (by using expenditure offsets) as “appropriation in aid.” In Australia, this responsibility is delegated to the executive branch: Section 31 of the *Financial Management and Accountability Act* directs the Finance Minister to determine whether government agencies are involved in business-like activity and the degree of discretion such departments and agencies ought to have over their own receipts. The Australian approach would only be advisable in most countries where the budget process is highly professional, apolitical, and transparent and where government officials are directly accountable for their policy decisions. Any approach that would allow the executive branch to approve expenditure offsets without the consent of the legislature would not be advisable in the context of the Russian budget process.

ACCOUNTABILITY, CONTROL, AND AUDITS All countries reviewed in this study have a firm line of executive accountability; managers of special funds are often directly accountable to the sectoral line minister within whose portfolio the fund is contained. That said, we should note that the international trend is toward increased flexibility for special fund managers. However, the trend coincides with the heightened ability of the center to effectively monitor and audit the financial transactions of individual government agencies and departments. In all countries reviewed, all budgetary funds and accounts are fully contained within the system of public accounts, and despite the special status of some accounts, all special account and budget funds are subject to essentially the same budget control, reporting, and audit measures as regular expenditure programs.

BUDGET REPORTING A final dimension of the use and accounting of extra-budgetary funds concerns the reporting of budgetary outcomes. All countries reviewed report on the main central government budget as well as providing additional information on the composition of their public sector. In this regard, the comprehensive nature of reporting in the Australian budget process stands out, although the United Kingdom is in the process of significantly expanding its reporting requirements also to present a “Whole of Government Account” and a “Whole of the Public Sector” report. Even in the United States, the budget specifically acknowledges the special nature of off-budget funds by reporting on-budget totals, off-budget totals, and overall aggregates.

Conclusions

The comprehensiveness of budget reporting in the countries reviewed here is no accident. Explicit consideration of the government’s primary fiscal balance is crucial in ensuring fiscal discipline and expenditure prioritization. At the same time, wider consideration of the government’s overall budget balance is needed to understand the macroeconomic impacts of government policies and to appreciate the overall need for external borrowing. Thus, for purposes of accountability and transparency, it would be advisable for the Russian Government to report on the performance and financial situation not only of its main budget accounts, but also of the entities that it controls, be they publicly owned enterprises or extra-budgetary funds. Although the existence of a small number of extra-budgetary funds could be justified at the federal level, there is no such justification at the subnational level. The Russian Federation should follow the example of Kazakhstan and other countries and abolish all extra-budgetary funds at the subnational level.

Appendix 8.2. Global Budgeting Trends and Their Relevance to Russia's Subnational Governments

A modern budget system should be able to perform three main functions: ensure that spending priorities legislated into a budget law correspond to the needs of the community, secure the conformity of actual expenditures with the legislated budget, and promote efficiency in service delivery through procedures that create incentives for minimizing costs and attaining the desired service level and quality.

Traditional budgets, as they are still used in the Russian Federation, are fundamentally organized by line department. This practice can easily lead to inconsistent spending priorities. For example, a government may wish to make reducing the incidence of sexually transmitted diseases, such as HIV/AIDS, a priority initiative. As a result, various departments and agencies may allocate funds for a variety of programs in support of this initiative. However, when a budget is organized by line departments, as opposed to program areas such as HIV/AIDS prevention, policymakers may fail to keep track of total annual expenditures allocated to a given program area across line departments and, over time, individual programs within the line departments may become duplicative and even develop conflicting goals.

Line-item budgeting, or budgeting by object of expenditure, is still a popular approach at the national and subnational levels worldwide owing to its simplicity and the control it allows. However, this budgeting approach has limited utility as a management tool. The greatest disadvantage of line-item budgeting is that there is no relation between expenditures and accomplishments, no concept of alternatives to policy choices, and no integration of planning, budgeting, and control. Line-item budgeting promotes inertia in that changes are generally only marginal adjustments from the previous year. Levels of service, organization structure, and methods of operation tend to become permanent, even if they are unsatisfactory.

In response to the shortcomings of the traditional line-item approach, a growing list of countries around the world has reformed their budget processes to achieve a more efficient resource allocation. These budgetary reforms have stressed the development of comprehensive budgets, where all budgetary expenditure proposals are considered at the same time and within the available fiscal envelope. The reforms are also replacing detailed ex ante control of expenditures with a more performance-oriented managerial climate. These changes are based on the belief that greater integration of performance measures into the budget process will improve resource allocation and overall performance of the public sector. It also reflects the view that concentrating on strategic issues such as reviewing program performance relative to clearly stated performance measures is more relevant for both macroeconomic and microeconomic efficiency than are detailed controls on inputs.⁹

Performance-based budgeting gained attention during the second half of the 1980s in countries including Australia, New Zealand, the United Kingdom, and the United States. These countries have continued to shift their budgeting paradigms from compliance to performance and at the same time have proceeded to give program administrators greater freedom to run their operations as they deem appropriate. Many forces shaped this transformation, including the adoption of management practices from the business sector; the persistent increase in the share of national income taken by the government; the ability of politicians to take credit for cutting expenditures without drawing blame for cutting programs; and the realization by budget officials that they were overcontrolling to the detriment of public-service provision.¹⁰ Most of these forces are or eventually will be in operation in the Russian Federation and as a result, there will be increasing pressure to adopt a performance-based paradigm for the budget process at all levels of government.

9. See, for example, Shand (1998).

10. See, for example, the discussion in Schick (1995).

A fundamental part of performance-oriented budget management is that program objectives must be clearly stated during the budgeting and evaluation periods. The achievement of these objectives must be clearly linked to their production requirements. In some cases performance targets are set out in formal contractual arrangements specifying the output or results that an organization or manager is committed to achieving with the agreed-upon level of resources.

Countries such as the United Kingdom, the United States, and others using performance-oriented techniques still use traditional measures for fiscal planning, such as the number of students in the case of education spending or the expected number of patients in the case of medical services. However, all of these countries have improved upon traditional methods by placing greater emphasis on fiscal planning in terms of the expected outcome, such as graduation rates, test scores, or access to health facilities. The most critical part of performance-oriented management is that outcomes and objectives need to be clear, measurable, and consistent with government policies.

A parallel development to the adoption of performance-oriented budget management has been the dismantling of the assessment controls that are related to inputs (needs or requirements to run government units) and that operate through the detailed itemization of different objects of expenditure. This latter trend also has been in response to the vast expansion of the government sector and the almost impossible task of creating different measures to assess expenditure needs for each single item across the great variety of governmental functions, agencies, and programs. Interestingly, this trend is fundamentally in the opposite direction of the continuous call in the Russian Federation for revitalizing the old Soviet budgeting approach of defining expenditure norms and controls for all kinds of subnational government expenditures and functions.

The international best practice shows that in the early stages of reforms and transformations of budgetary practices the emphasis should be on less complex budget procedures, the outcomes achieved by government spending units, and the evaluation of managers and personnel according to these outcomes. Sophistication will come with practice and with the development of better performance measures and outcomes over time. The increase in budget flexibility and discretion in the allocation of public resources should come only after the mechanisms are in place for holding managers accountable.

9

Conclusion

The first decade of the Russian Federation—and its transition toward a new federalist arrangement—was marked by the escalation and eventual winding down of centrifugal forces on the country that were similar to those that had led to the disintegration of the Soviet Union. This backdrop determined the shape and form that the Russian federal structure would assume during the transition. Although the formal acceptance of asymmetric relations in the 1993 Constitution diminished the centrifugal forces and brought some predictability and order to the system of intergovernmental fiscal relations, this “bargaining federalism” has not been without costs in the Russian Federation. Besides generating mistrust and resentment among the regions, the asymmetric approach to federalism in general and the practice of the bilateral treaties between the federal government and regions in particular has had other significant costs, including mounting fiscal pressure on the federal budget, ultimately culminating in the financial crisis of 1998.

During the 1990s, following its practice of securing checks and balances, the Yeltsin administration lobbied for constitutional provisions of local self-government, thus trying to challenge the power of the then-mighty regional governors. Influenced by the Council of Europe principles (which Russia joined in 1996), federal legislation defines local self-government as the management of local affairs by local residents under their own responsibility. However, as a sign of the power of the regions during that period, regional governments had absolute discretion in the design and organization of their relationships with their local governments. The establishment of fundamentally decentralized institutions, local elections, decentralization of resources between regional and local governments, and established rules of financial accounting and budgeting have all failed so far to lead to an accountable fiscal system at the regional-local level. In part this is due to the inappropriate intergovernmental arrangements prevailing in most of the regions. The competences of local governments are not clearly defined and their lack of revenue-raising powers makes them dependent on the revenue decisions of regional authorities. As a result, citizens are reluctant to engage in local budgetary issues. It is as if they are saying, “If local governments are not delegated enough fiscal autonomy, it is not worth spending time on monitoring local officials.” Thus, while local government bodies might have some autonomy, the current situation hardly reflects the principle in the Council of Europe’s charter of “local residents managing their affairs.”

One clear trend in recent federal policy is the new willingness of the federal government to intervene in the fiscal arrangements between regional governments and their local governments. An important recent development in this direction has been the appointment by President Putin of a high-level commission in 2001 (known as the Kozak Commission) for redefining and reforming intergovernmental fiscal relations in Russia. The regional-local sector is thoroughly covered in the recommendations of the Kozak Commission. Among other things the findings include the need for a clear separation of competences among the different levels of government. We summarize some of the Kozak proposals below.

Local government structure. The Kozak proposal effectively mandates a uniform two-tier structure of local government in the entire territory of the Russian Federation. This essentially updates federal legislation to match the local government structure observed in Russian regions. This improvement is expected to eliminate the current conflict between the centralization of tax policy powers at the federal level and the decentralization to the regional level of the

discretion over local government structure and functions. Theoretically, there were two possible resolutions to this conflict: centralization of authority over local government structure at the federal level, and the delegation to the regional level of the authority over the delineation of subnational revenue-raising powers between the regional and local levels. Apparently, the federal government opted for the first option, which has several advantages. It may make it somewhat easier for the federal Ministry of Finance to control the finances of local governments from Moscow. Under this option, revenue sources can be allocated more efficiently by assigning more unevenly distributed revenue sources to the higher tier of local government and more uniformly distributed revenue sources to the lower tier of local government. However, subnational governments will continue to be deprived of an adequate level of revenue autonomy, which will hurt accountability and the overall efficiency of the government sector.

It is important to put the Kozak proposal in perspective and recognize that there is no powerful need for one single approach toward local governments to be applied in all regions in Russia. Perhaps a more practical solution would allow regional governments to choose from a number of approaches to local government structure. There is a tremendous difference in terms of local government capacity between urban and rural administrative units, especially at the lowest level of government. The strict imposition of regional-local delineation of responsibilities that allow for no distinctions between urban and rural settlements might therefore become a source of rigidity. It is true that the Kozak proposal will allow regions to treat municipalities according to basic factors such as rural versus urban character and population size when needed—but so could a more flexible proposal that would allow the regions to adopt a model more consonant with their needs and even traditions.

Expenditure assignments. On the expenditure side, the Kozak proposals represent a significant improvement over the ambiguous delineation of responsibilities outlined in the Constitution and more recently in the *Budget Code*. It is significant that the proposed legislation enumerates exclusive responsibilities for subnational governments other than just managing their own operations and property. Besides increasing clarity in the assignment of responsibilities, the Kozak proposals improve the delineation of expenditure responsibilities in other respects. First, they make clear the principle that for “delegated” responsibilities the upper-level government delegating those responsibilities has the power to regulate but also the obligation to ensure there is sufficient funding. Second, they make clear that for all other local responsibilities (own or non-mandated), upper level governments cannot introduce expenditure norms.

Several problems should be expected from the implementation of the Kozak proposals. Allowing district governments to have administrative responsibility over municipalities at the settlement level in an environment where mutual trust is lacking would likely result in instances where districts would attempt to abuse their control over the lower-level local governments in contradiction with regional law. Experiences in transitional and developing economies with multilevel governments show that each intermediate link weakens the government structure, allowing money to leak out of the system and reducing the overall accountability of the public sector.

While being too rigid in certain respects, such as the structure and functions of local government, the proposed arrangements might be too ambiguous regarding others. We have seen that for concurrent or shared functions, the Kozak proposals fall short of explicitly specifying the “attributes” for each shared responsibility among two or more levels of government. This ambiguity only adds to the daunting task of public administration reform that the federal government is currently facing. Functional responsibilities can be further clarified by specifying in the law what level of government is responsible for regulating, financing, and implementing each of the shared responsibilities. This should also help eliminate the fuzzy terms still used in the Kozak proposals. However, there is no guarantee that the assignment of specific attributes (on regulation, financing, or delivery) for the existing programs would completely clarify responsibilities. For example, it is possible for regulation powers to be attributed to two levels of gov-

ernment. In these cases, there should be further clarification in the law on what particular power of regulation the different levels of government will have.

There is no formally legislated expenditure assignment that can be exhaustively complete. There will always be important details as well as new situations that will need to get resolved in practice. Solving the differences in perspective of the two levels of government should not be a one-way street, with the upper-level governments always dictating solutions. The effectiveness of expenditure assignments and the day-to-day operations of the intergovernmental system could be much improved by creating or further enhancing institutions and mechanisms for coordination and dialogue among different levels of government in all sorts of expenditure-assignment issues. Those responsible for general education at the regional and local levels would meet periodically and have open channels of communication for all matters of interpretation and uncertainty in their concurrent responsibilities. There are also significant opportunities for cooperation among local government units for joint provision of services (through public or private companies, or an association among them) in order to explore economies of scale (by merging or associating existing companies) and fully consider existing externalities among jurisdictions.

Revenue assignments. Unlike the 1997 *Law on the Financial Foundations of Local Self-Government in the Russian Federation*, the Kozak laws lack any provisions securing greater vertical fiscal balance among different levels of government. The 1997 law required that regional governments should, on average, meet some minimum rates for sharing revenue from major taxes with local governments. By contrast, the proposed law does not separate vertical balancing from horizontal equalization and in fact allows for the allocation of the equalization grants through tax-revenue retention. This creates a danger that all tax-revenue sharing will be treated as equalization grants rather than long-term funding entitlements. In fact, since 2000 the *Budget Code* has been requiring tax-sharing rates to be stable for a period of no less than three years. Nevertheless, most regions still treat revenue sharing above the minimum requirements as equalization grants and alter retention rates on an annual basis. To secure vertical balance at the local level, federal legislation might have to require long-term assignments of tax-revenue shares to cover a certain portion of consolidated local expenditures. This would be less restrictive than the currently mandated minimum rates for sharing revenue from particular taxes. The unevenness of fiscal capacity may pose some difficulties in the determination of revenue shares. However, negative transfers would allow regional authorities to deal with outliers resulting from uniform tax-revenue sharing.

As is well known, the only flexible way to reach vertical balance is through providing all levels of government with sufficient revenue autonomy, at the subnational level most preferably in the form of discretion over tax rates for some productive tax bases. These decisions should be part of long-term institutional arrangements and related to the established expenditure obligations of local governments. It is commendable that in the current phase of reforms, unlike in the early 1990s, changes to the assignment of revenue sources are to follow the legislated delineation of functions and not the other way around. However, even before the split of expenditure responsibilities was finalized, one could easily have predicted a problem with the lack of subnational revenue autonomy to arise unless productive tax instruments were provided to subnational authorities. Unfortunately, while the proposed tax reforms eliminate a number of important (revenue-wise) tax instruments that were available to subnational governments before, they offer no new tax instruments to subnational governments.

Overall, it appears that the federal government underestimates the importance of subnational governments' revenue autonomy for the well-being of the nation. Although the Kozak Commission initially declared the importance of subnational revenue autonomy, in the course of its work this principle fell prey to "federal interests." For example, the Commission accepted the evidently false argument that surcharges on the personal income tax (PIT) would ruin the benefits of a flat-rate tax, such as reducing tax evasion. Moreover, the Commission justified the

monopolization of tax policy by the federal government for the sake of uniform treatment of taxpayers and a uniform environment for economic activities. While these goals appear valid they are not necessarily incompatible with giving larger revenue autonomy to subnational authorities.

Tax-treatment uniformity concerns the legal interpretation and geographical equality in the selection of tax accounts for audit. Uniform application of the tax laws across taxpayers demonstrates that the tax administration is committed to the fair treatment of taxpayers. Undoubtedly, a well-publicized emphasis on uniform treatment enhances the perception of the tax system as fair and increases voluntary compliance.¹ However, administrative uniformity would not be hurt if subnational authorities introduced surtaxes on uniform tax bases across the country defined by federal taxes and piggybacked on the federal tax administration. Under the centralized administration, taxpayers continue to receive uniform treatment in assessment, audit, penalties, and appeals, irrespective of where they reside. In fact, that uniformity was demonstrated in Russia with the successful piggybacking of regional authorities on the federal tax on enterprise profits.

The uniformity with respect to the tax burden is a more questionable argument. It might appear that tax holidays on the profit tax provided by some Russian regions (before being finally outlawed) did constitute a “rob thy neighbor” strategy. But some problems resulted from the initial violation of the benefit principle in making the profit tax a predominantly subnational revenue source. The bottleneck that international experience has revealed in administering a subnational corporate income tax (CIT) has been the establishment of a practical link between the amount of CIT paid in each jurisdiction and the benefits from subnational government services enjoyed by the taxpayer. If a non-benefit tax is applied to a mobile base, such as book profits, inefficiencies can arise, as taxpayers would try to reduce their tax liability without losing their benefits from public provision. However, if the benefit principle were preserved, the mobility of a tax base would not bring about distortions but increased efficiency through arbitraging tax/benefit differentials. Thus, it is not the level of taxation that concerns economic agents but the tax/benefit differentials. The uniformity of the latter is better achieved with healthy competition among jurisdictions than with hierarchical constraints.

Intergovernmental grants. Another source of ambiguity in the Kozak proposals is the lack of requirements for the methodology of distributing equalization grants at the regional-local level. The proposed laws refer to the *Budget Code*, which currently requires only the fulfillment of “social standards” in distributing transfers. This constitutes a regression from the 1997 *Law on the Financial Foundations of Local Self-Government in the Russian Federation*, which required usage of a stable formula and provided a list (that in some ways was too specific) of parameters to be used. We must note that these requirements were rarely met in practice. Nevertheless, it would seem desirable for the federal government to introduce more specific requirements for the equalization methodology than those currently provided in the *Budget Code*, though perhaps not as detailed as in the *Law on the Financial Foundations of Local Self-Government in the Russian Federation*. It is necessary to find a finer balance between the flexibility of regional governments and minimum standards of sound fiscal practice. Although certain parameters of the equalization scheme can be determined annually, the general framework for grant allocation should be a long-term decision. For example, the base level for expenditure standards can be set each year based on macroeconomic conditions; however, the methodology for assessing the cross-jurisdictional cost differentials should be fixed on a long-term basis.

All in all, the recommendations of the Kozak Commission should result in an overall improvement of intergovernmental relations at the regional-local level in Russia. However,

1. See, for example, the discussion in Vehorn and Ahmad (1997).

more work will be needed in the near future to further fine-tune the Commission's recommendations. Further evolution of the system of regional-local relations should be expected as inter-governmental fiscal relations are a continually evolving process—one that will continue to demand additional study and reform for years to come. Thus, after this book was written news came out that President Putin plans to start appointing regional governors, with the approval of regional legislatures. While these measures represent a blow to the foundations of Russian federalism, they also raise the importance of local government as a vehicle of self-government and a "school of democracy." We hope that this study facilitates and contributes to the ongoing efforts to improve regional-local finance in the Russian Federation.

References

The word *processed* describes informally reproduced works that may not be commonly available through libraries.

- Alexeev, Michael, and Galina Kurlyandskaya. 2003. "Fiscal Federalism and Incentives in a Russian Region." *Journal of Comparative Economics* 31: 20–33.
- Alfandari, Gilles, and Mark E. Schaffer. 1996. "'Arrears' in the Russian Enterprise Sector." Centre for Economic Reform and Transformation Discussion Paper DP96/08. CERT, Edinburgh.
- Bahl, Roy. 1999. *Fiscal Policy in China: Taxation and Intergovernmental Relations*. San Francisco, Calif.: The 1990 Institute.
- Bahl, Roy, and the Intergovernmental Fiscal Relations Team. 1999. "Intergovernmental Fiscal Relations in Leningrad Region." International Studies Working Paper #99-2, Andrew Young School of Policy Studies, Georgia State University, Atlanta.
- Bahl, Roy, and Johannes F. Linn. 1992. *Urban Public Finance in Developing Countries*. New York: Oxford University Press.
- Bahl, Roy, and Jorge Martinez-Vazquez. Forthcoming. "Fiscal Federalism and Economic Reform in China." In T. S. Srinivasan and Jessica S. Wallach, eds., *Fiscal Federalism in a Global Environment*. Cambridge, New York, Melbourne: Cambridge University Press.
- Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy* 87(5): 940–71.
- Bayoumi, Tamim, Morris Goldstein, and Geoffrey Woglom. 1995. "Do Credit Markets Discipline Sovereign Borrowers? Evidence from U.S. States." *Journal of Money, Credit and Banking* 27(4): 1046–59.
- Bird, Richard M. 1999. "Rethinking Subnational Taxes: A New Look at Tax Assignment." International Monetary Fund Working Paper WP/99/165. Washington, D.C.
- Bird, Richard M., and Pierre-Pascal Gendron. 2001. "VATS in Federal States: International Experience and Emerging Possibilities." International Studies Working Paper #01-4, Andrew Young School of Policy Studies, Georgia State University, Atlanta.
- Bird, Richard M., and Enid Slack. 2002. "Land and Property Taxation Around the World: A Review." *Journal of Property Tax Assessment and Administration* 7(3): 31-80.
- Blanchard, Olivier, and Andrei Shleifer. 2000. "Federalism with and without Political Centralization: China versus Russia." Working Paper, Massachusetts Institute of Technology, Cambridge.
- Break, George F. 1980. *Financing Government in a Federal System*. Washington, D.C.: Brookings.
- Buchanan, James M. 1965. "An Economic Theory of Clubs." *Econometrica* 32: 1–14.
- Capeci, John. 1994. "Local Fiscal Policies, Default Risk, and Municipal Borrowing Costs." *Journal of Public Economics* 53(1): 73–89.
- Commander, Simon, and Mark Schankerman. 1997. "Enterprise Restructuring and Social Benefits." European Bank for Reconstruction and Development (EBRD) Working Paper No. 22, London.
- Commander, S., I. Dolinskaya, and C. Mumssen. 2000. "Determinants of Barter in Russia: An Empirical Analysis." IMF Working Paper WP/00/155, Washington, D.C.

- Colclough, Christopher, ed. 1997. *Public Sector Pay and Adjustment: Lessons from Five Countries*. London and New York: Routledge.
- Council of Europe. 1994. "Draft Report on Monitoring the Implementation of the European Charter of Local Self-Government. Strasbourg.
- Crédit local de France and Crédit communal de Belgique. 1997. *Local Finance in the Fifteen Countries of the European Union*. Brussels: Dexia.
- de Figueiredo, Rui J. P., and Barry R. Weingast. 2001. "Constructing Self-Enforcing Federalism in the Early United States and Modern Russia." January, processed.
- Diamond, Jack. 2002. "Budget System Reform in Transitional Economies: The Experience of Russia." International Monetary Fund Policy Working Paper: WP/02/22, Washington, D.C.
- Ebel, Robert, and Serdar Yilmaz. 2002. "On the Measurement and Impact of Fiscal Decentralization." Policy Research Working Paper Series no. WPS 2809. World Bank, Washington, D.C.
- Ebrill, Liam, Michael Keen, Jean-Paul Bodin, and Victoria Summers. 2001. *The Modern VAT*. Washington, D.C.: International Monetary Fund.
- Elliott, Robert F., and Keith A. Bender. 1997. "Decentralization and Pay Reform in Central Government: A Study of Three Countries." *British Journal of Industrial Relations* 35(3): 447–75.
- Falzon, Micheline, Marie-Pierre Peretti-Stahl, and Axelle Verdier. 2002. *Local Finance in the Fifteen Countries of the European Union*. Paris: Dexia.
- Firestone, Allan, and Tax Administration Team. 1998. "Modernization of the Tax Administration of the Russian Federation." Georgia State University Working Paper, ISP WP 98-6, Atlanta.
- Freinkman, Lev, Daniel Treisman, and Stepan Titov. 1999. "Subnational Budgeting in Russia: Preempting a Potential Crisis." Technical Paper No. 452, World Bank, Washington, D.C.
- Goskomstat. 2001. *Russian Statistical Yearbook*. Moscow.
- Greene, William. 1993. "The Econometric Approach to Efficiency Analysis." In Harold O. Fried, C. A. Knox Lovell, and Shelton S. Schmidt, eds., *The Measurement of Productive Efficiency: Techniques and Applications*. New York: Oxford University Press.
- Högye, Mihály, Hugo Van Hassel, Gunar Olsson, Francisca Sabbe, and Katarina Ott. 2000. "Comparative Approaches to Central and Eastern European Tax Reform." In Mihály Högye, ed., *Local and Regional Tax Administration in Transition Countries*. Budapest, Hungary: The Local Government and Public Service Reform Initiative.
- Horváth, Tamás M. ed. 2000. *Decentralization: Experiments and Reforms*. Budapest, Hungary: The Local Government and Public Service Reform Initiative.
- Hughes, Michael, Michael Clarke, Hubert Allen, and Declan Hall. 1998. "The Constitutional Status of Local Government in Other Countries." Report Prepared for the Commission on Local Government at the Scottish Parliament. The Scottish Office Central Research Unit. Edinburgh.
- Katz, Ellis. 1999. "Local Self-Government in the United States." *Issues of Democracy: An Electronic Journal of the U.S. Information Agency*, Vol. 4, No. 1.
- Keen, Michael, and Stan Smith. 2000. "Viva VIVAT!" *International Tax and Public Finance*, Vol. 7: 741–51.
- Kourliandskaia, Galina, Yelena Nikolayenko, and Natalia Golovanova. 2002. "Local Government in the Russian Federation." In Victor Popa and Igor Munteanu, eds., *Developing New Rules in the Old Environment. Local Governments in Eastern Europe, Caucasus and Central Asia*. Budapest, Hungary: The Local Government and Public Service Reform Initiative.

- Kuzmin, A. I., E. A. Kachanova, I. A. Pykhova, S. S. Pushkarve, and D. V. Osintsev. 2002. "Local Government Budgeting—Russia." In Mihály Högye, ed., *Local Government Budgeting*. Budapest, Hungary: The Local Government and Public Service Reform Initiative.
- Lane, Timothy D. 1993. "Market Discipline." *International Monetary Fund Staff Papers* 40(1): 53–88.
- Lavrov, Alexei M., John M. Litwack, and Douglas Sutherland. 2001. "Fiscal Federalist Relations in Russia: A Case for Subnational Autonomy." OECD, processed.
- Lykova, Lyudmila, Natalia Krylova, and Alla Chebanova. 1999. "Public Debt of Subjects of the Russian Federation: Current State and Perspectives." *Federalism* 1999(3) (in Russian).
- Malme, Jane H., and Natalia Kalinina. 2001. "Property Tax Developments in the Russian Federation." In Malme, H. Jane, and Joan M. Youngman, eds., *The Development of Property Taxation in Economies in Transition: Case Studies from Central and Eastern Europe*. WBI Learning Resources Series. Washington, D.C.: World Bank.
- Martinez-Vazquez, Jorge. Forthcoming. "Asymmetric Federalism in Russia: Cure or Poison?" In R. Bird, R. Ebel, and C Wallich, eds., *Asymmetric Federalism*. Cheltenham, U.K. and Northampton, Mass.: Edward Elgar.
- Martinez-Vazquez, Jorge, and Christine Wong. 2003. "Two Large Experiments in Fiscal Decentralization Compared: China and Russia." International Studies Program Working Paper 03-13, Andrew Young School of Policy Studies, Georgia State University, Atlanta.
- Martinez-Vazquez, Jorge, and L. F. Jameson Boex. 2001. *Russia's Transition to a New Federalism*. WBI Learning Resources Series. Washington, D.C.: World Bank.
- Martinez-Vazquez, Jorge, and Robert M. McNab. 2000. "The Tax Reform Experiment in Transitional Countries." *National Tax Journal* 13(2): 273–98.
- Martinot, E. 1997. *Investments to Improve the Energy Efficiency of Existing Residential Buildings in Countries of the Former Soviet Union*. Studies of Economies in Transformation, no. 24. Washington, D.C.: World Bank.
- Martins, M. R. 1995. "Size of Municipalities, Efficiency and Citizen Participation: A Cross-European Perspective." *Environment and Planning C: Government and Policy* 13: 441–58.
- McLure, Charles. 1999. "The Tax Assignment Problem: Conceptual and Administrative Considerations in Achieving Subnational Fiscal Autonomy." Paper prepared for the core course on Intergovernmental Relations and Local Financial Management, World Bank Institute, Washington, D.C.
- Mikesell, John L. 1999. *Fiscal Administration: Analysis and Applications for the Public Sector*. Fort Worth, Tex.: Harcourt Brace College Publishers.
- Mouritzen, Poul Erik. 1989. "City Size and Citizen Satisfaction: Two Competing Theories Revisited." *European Journal of Political Research* 17: 661–88.
- Musgrave, Richard A. 1998. "Who Should Tax, Where and What?" In Wallace E. Oates, ed., *The Economics of Fiscal Federalism and Local Finance*. Cheltenham, U.K. and Northampton, Mass.: Edward Elgar.
- Musgrave, Richard A., and Peggy B. Musgrave. 1976. *Public Finance in Theory and Practice*. New York: McGraw-Hill.
- Norton, Alan. 1994. *International Handbook of Local and Regional Government: A Comparative Analysis of Advanced Democracies*. Aldershot, Hants, U.K. and Brookfield, Vt.: Edward Elgar.

- Nunberg, Barbara. 1999. "Transforming the Prototype: A New Administrative Order for Post-Soviet Russia?" In Barbara Nunberg, Luca Barbone, and Hans-Ulrich Derlien, eds., *The State after Communism: Administrative Transitions in Central and Eastern Europe*, pp. 155–94. Regional and Sectoral Studies. Washington, D.C.: World Bank.
- Oates, Wallace E. 1972. *Fiscal Federalism*. New York: Harcourt Brace Jovanovich.
- OECD. 1997. *OECD Economic Surveys: Russian Federation*. Paris.
- . 1999. *Revenue Statistics: 1965–1998: Special Features, Taxing Powers of State and Local Government, the Interpretation of Tax-to-GDP Ratios, the Impact of GDP Revisions on Reported Tax Levels*. Paris.
- . 2000. *OECD Economic Surveys: Russian Federation*. Paris.
- Potapchik, Elena G. 2002. "Review of the Russian Legislation in the Sphere of Health Insurance." *Governance of Social Security: Social Insurance, Medical Insurance and Pensions*. Moscow: TACIS.
- RET. 1997. "1997, Q3." *Russian Economic Trends Quarterly*. Russian European Centre for Economic Policy, Moscow.
- Rodden, Jonathan A., and Gunnar S. Eskeland, eds. 2003. *Fiscal Decentralization and the Challenge of Hard Budget Constraints*. Cambridge: Massachusetts Institute of Technology.
- Rubinfield, Daniel L. 1983. "Tax Assignment and Revenue Sharing in the United States." In Charles E. McLure, Jr., ed., *Tax Assignment in Federal Countries*. Canberra: Australian National University.
- Schiavo-Campo, Salvatore, and Daniel Tommasi. 1999. *Managing Government Expenditure*. Manila, Philippines: Asian Development Bank.
- Schick, Allen. 2001. "Does Budgeting Have a Future?" Prepared for the 22nd Annual Meeting of Senior Budget Officials, Paris, May 21–22, 2001, OECD.
- Shand, David. 1998. "Budgetary Reforms in OECD Member Countries." *Journal of Public Budgeting, Accounting and Financial Management* 10(Spring): 63–88.
- Sharpe, L. J. 1970. "Theories and Values of Local Government." *Political Studies* 18: 153–74.
- Shleifer, Andrei, and Daniel Treisman. 2000. *Without a Map: Political Tactics and Economic Reform in Russia*. Cambridge: Massachusetts Institute of Technology.
- Silvestre J. J., and F. Eyraud, eds. 1995. "Pay Determination in the Public Sector: An International Comparison between France, Great Britain and Italy." Labour Law and Labour Relations Programme Occasional Paper, No. 6. International Labour Office, Geneva.
- Stewart, Kitty. 1997. "Are Intergovernmental Transfers in Russia Equalizing?" EUI Working Papers, ECO No. 97/22. European University Institute, Florence, Italy.
- Swianiewicz, Pawel. 2003. "Foundations of Fiscal Decentralization, Benchmarking Guide for Countries in Transition." LGI Discussion Paper No. 26. The Local Government and Public Service Reform Initiative, Budapest, Hungary.
- Ter-Minassian, Teresa, ed. 1997. *Fiscal Federalism in Theory and Practice*. Washington, D.C.: International Monetary Fund.
- Ter-Minassian, Teresa, and John Craig. 1997. "Control of Subnational Government Borrowing." In Teresa Ter-Minassian, ed., *Fiscal Federalism in Theory and Practice*. Washington, D.C.: International Monetary Fund.
- Tiebout, Charles. 1956. "A Pure Theory of Local Public Expenditures." *Journal of Public Economy* 64: 416–24.

- Timofeev, Andrey. 2003. "Determinants of Decentralization within Russian Regions." LGI Discussion Paper No. 23. The Local Government and Public Service Reform Initiative, Budapest, Hungary.
- . 2004a. "Land and Property Taxes in Russia." In Richard M. Bird and Enid Slack, eds., *International Handbook of Land and Property Taxation*. Cheltenham, U.K. and Northampton, Mass.: Edward Elgar.
- . 2004b. *Fiscal Decentralization in the Context of Transition*. Ph.D. thesis. Prague, Czech Rep.: Center for Economic Research and Graduate Education of Charles University.
- Tolkushkin, A. V. 2001. *History of Taxation in Russia*. Moscow: Yurist (in Russian).
- Treisman, Daniel. 1996. "The Politics of Intergovernmental Transfers in Post-Soviet Russia." *British Journal of Political Science* 26(2): 299–335.
- United States General Accounting Office. 1994. *Tax Administration: State Tax Administrators' Views on Delinquent Tax Collection Methods*. GAO/GGD-94-59FS, Washington, D.C., February.
- United States Government Technical Assistance Team. 1998. "The Choice between the VAT and the Retail Sales Tax in the Russian Federation." Georgia State University, Andrew Young School of Policy Studies, Atlanta, Ga.
- Vedi. 1998. *General Overview of Municipal Bonds*. Moscow: Econometric Unit—"Vedi."
- Vehorn, Charles L., and Ehtisham Ahmad. 1997. "Tax Administration." In Teresa Ter-Minassian, ed., *Fiscal Federalism in Theory and Practice*. Washington, D.C.: International Monetary Fund.
- Vinokur, Anatoly, Joana Godinko, Christopher Dye, and Nico Nagelkerke. 2001. "The TB and HIV/AIDS Epidemics in the Russian Federation." Technical Paper no. 510, World Bank, Washington, D.C.
- von Hagen, Jürgen, and Barry Eichengreen. 1996. "Federalism, Fiscal Restraints, and European Monetary Union." *American Economic Review* 86(2): 134–38.
- Walberg Peder, Martin McKee, Vladimir Shkolnikov, Laurent Chenet, and David A. Leon. 1998. "Economic Change, Crime, and Mortality Crisis in Russia: A Regional Analysis." *British Medical Journal* 317: 312–18.
- Wetzel, Deborah. 2002. *Decentralization in the Transition Economies: Challenges and the Road Ahead*. Washington, D.C.: World Bank.
- World Bank. 1996. *Fiscal management in Russia*. Washington, D.C.
- . 2001. *Russian Federation Fiscal Federalism and Regional Fiscal Reform Project*. Washington, D.C., July.
- . 2002a. *Lithuania—Issues in Municipal Finance*. Sector Report No. 23716. Washington, D.C.
- . 2002b. *China—National Development and Sub-National Finance: A Review of Provincial Expenditures*. Report No. 22951-CHA. Washington, D.C.
- . 2002c. *Transition—The First Ten Years: Analysis and Lessons for Eastern Europe and the Former Soviet Union*. Washington, D.C.
- Zamiatina M., B. Beskrovnaia, T. Vinogradova, and D. Torhov. 2002. *Evaluation of Transparency and Public Participation in the Budget Process*. Report. St. Petersburg: Strategy.
- Zhuravskaya, E. V. 2000. "Incentives to Provide Local Public Goods: Fiscal Federalism, Russian Style." *Journal of Public Economics* 76: 337–68.

WORLD BANK INSTITUTE

The World Bank Institute (WBI) develops capacity in World Bank client countries through courses and seminars, policy advice, diagnostic tools, and other products and services to help countries achieve their development goals.

Through traditional face-to-face and distance learning methods, WBI and its partners around the world work with policymakers, technical experts, business and community leaders, and civil society stakeholders to foster the kinds of analytical and networking skills that support effective socioeconomic programs and policymaking. The Institute collaborates with World Bank operations staff on high priority themes including human development, poverty reduction and economic management, environmentally and socially sustainable development, and finance and private sector development.



THE WORLD BANK



0-8213-6557-6