Institutional Elements of Tax Design and Reform
Recent World Bank Technical Papers


No. 429 Gary McMahon, José Luis Zveva, Alberto Pasco-Font, and José Miguel Sánchez, An Environmental Study of Artesian, Small, and Medium Mining in Bolivia, Chile, and Peru

No. 430 Maria Dakilakis, Court Performance around the World: A Comparative Perspective

No. 431 Severin Koddereit, Reforms in Albanian Agriculture: Assessing a Sector in Transition

No. 432 Luiz Gabriel Azevedo, Musa Asad, and Larry D. Simpson, Management of Water Resources: Bulk Water Pricing in Brazil

No. 433 Malcolm Rowat and José Astigarraga, Latin American Irrigation Systems: A Comparative Assessment


No. 435 Lain Begg, EU Investment Grants Review

No. 436 Roy Prosterman and Tim Harstad, ed., Legal Impediments to Effective Rural Land Relations in Eastern Europe and Central Asia: A Comparative Perspective

No. 437 Csaba Csaki, Michel Debatusse, and Oskar Lionisch, Food and Agriculture in the Czech Republic: From a "Velvet" Transition to the Challenges of EU Accession


No. 439 Mustapha Nabi, Financial Integration, Vulnerabilities to Crisis, and EU Accession in Five Central European Countries

No. 440 Robert Bruce, Ioannis Kassides, and Lothar Kneifel, Overcoming Obstacles to Liberalization of the Telecommunications Sector in Estonia, Poland, the Czech Republic, Slovenia, and Hungary: An Overview of Key Policy Concerns and Potential Initiatives to Facilitate the Transition Process

No. 441 Bartlomiej Kamiński, Hungary: Foreign Trade Issues in the Context of Accession to the EU

No. 442 Bartlomiej Kamiński, The Role of Foreign Direct Investment and Trade Policy in Poland's Accession to the European Union

No. 443 Luc Leclui, John Elder, Christian Hirtado, François Kantra, Kamal Siblini, and Maurizia Tovo, DeMISifying MIS: Guidelines for Management Information Systems in Small Firms

No. 444 Robert P. Townsend, Agricultural Incentives in Sub-Saharan Africa: Policy Challenges

No. 445 Tim Hill, Forest Management in Nepal: Economics of Eucalyptus

No. 446 Gordon Hughes and Magda Lovei, Economic Reform and Environmental Performance in Transition Economies

No. 447 R. Maria Salch and Ariid Dinar, Evaluating Water Institutions and Water Sector Performance

No. 448 Keith Obliquis and J. Raymond Peter in association with Guatem Pung, Halka M. Qaddumi, and Jayantha Perera, Transferring Irrigation Management to Farmers in Audra Pradesh, India


No. 450 World Bank, Privatization of the Power and Natural Gas Industries in Hungary and Kazakhstan

No. 451 Lev Freinkman, Daniel Treisman, and Stephen Tito, Subnational Budgeting in Russia: Preempting a Potential Crisis

No. 452 Bartlomiej Kamiński and Michelle Riboud, Foreign Investment and Restructuring: The Evidence from Hungary

No. 453 Gordon Hughes and Julia Bucknall, Poland: Complying with EU Environmental Legislation

No. 454 Dale F. Gray, Assessment of Corporate Sector Value and Vulnerability: Links to Exchange Rate and Financial Crises


No. 456 Mary Canning, Peter Moock, and Timothy Helenick, Reforming Education in the Regions of Russia

No. 457 John Gray, Kazakhstan: A Review of Farm Restructuring

No. 458 Zvi Lerman and Csaba Csaki, Ukraine: Review of Farm Restructuring Experiences

No. 459 Gloria La Cava and Rafaela Y. Nanetti, Albania: Filling the Vulnerability Gap

No. 460 Ayse Kudat, Stan Peabody, and Caglar Keyder, eds., Social Assessment and Agricultural Reform in Central Asia and Turkey

No. 461 T. Rand, J. Haukohl, and U. Marxen, Municipal Solid Waste Incineration: Requirements for a Successful Project

(List continues on the inside back cover)
Institutional Elements of Tax Design and Reform

Edited by
John McLaren

The World Bank
Washington, D.C.
# Table of Contents

Foreword iv  
Abstract v  
Acknowledgments vi  
Contributors vii  

1. Overview 1  

2. Corruption in Tanzania: Counting and Franchise Bidding  
   Dan Kobb 11  

3. Value Added Tax Evasion, Auditing and Transactions Matching  
   Ira Gang and Arindam Das-Gupta 25  

4. Fiscal Pathologies and Federal Politics: Understanding Tax Arrears in Russia’s Regions  
   Daniel Treisman 49  

5. The Political Economy of Center-State Fiscal Transfers in India  
   M. Govinda Rao and Nirvikar Singh 69  

6. Corruption and the Organization of Tax Administration: Non-Smithian Benefits from Specialization  
   John McLaren 125  

7. Incentive Effects of Performance-Based Rewards in Tax Administration  
   Emilson C. D. Silva 135  

8. Leadership or Institutions? Reputations, Retaliation and Opportunistic Taxation  
   Margaret S. McMillan 153
Foreword

Reform of the tax system continues to be a major element of policy reform in many developing countries. While many lessons can be learned from past reform efforts and substantial technical assistance for tax reform is being provided, the challenge to successfully complete a reform initiative remains substantial. Many countries struggle with the difficult task to coordinate policy reforms with administrative and institutional reforms, and further analysis of specific reform approaches and experience and of the links between tax policy and tax administration reform are needed to facilitate tax reform in developing countries.

In this light, a joint Columbia University/World Bank conference entitled “Institutional Elements of Tax Design and Reform” was held at Columbia University on February 18-19, 2000. The conference was intended to promote discussion among tax economists and practitioners on the links between tax policy and tax administration reform. The discussion focused on the institutional constraints on tax systems in the context of developing countries, where tax policy is generally complex, and administration capacity is notably weak.

The World Bank has long been active in supporting the strengthening of revenue systems in developing countries to raise the necessary funds for pro-poor expenditure programs, improve good governance and private sector development, and reduce corruption. The publication of a selection of papers presented at the aforementioned conference and revised following the discussions and comments received, intends to contribute to the on-going academic and practical exchange of ideas on tax policy design and reform in developing countries. In these papers, the authors provide thought-provoking perspectives on the way tax reforms are formulated and implemented. The book has been sponsored by the Public Sector Group of the Poverty Reduction and Economic Management (PREM) Network of the World Bank and edited by John McLaren, Department of Economics, University of Virginia.

Michael Engelschalk
Head, Tax Policy and Administration Thematic Group
Public Sector Group
Poverty Reduction and Economic Management Network
Abstract

This is a collection of papers that study the constraints on fiscal systems imposed by problems of institutions, administration, and incentives in developing and post-Socialist economies. Two chapters focus on the administration of indirect taxation. The chapter by Kobb is a case study of indirect taxation in Tanzania. It shows how evasion can be documented and quantified, and studies a particular type of reform aimed at curbing evasion: Franchising, or privatising the right to tax, which has been tried in several Tanzanian towns as a way of collecting vendor fees for access to a public market. The chapter by Gang and Das-Gupta is a theoretical study of evasion under a value-added tax (VAT), and the inefficiencies that it can create. Two chapters examine case studies in fiscal federalism in which the determination of fiscal outcomes is, at least to a considerable degree, a matter of bargaining or gamesmanship between political entities in the center and in the periphery. The paper by Treisman studies Russia, and the paper by Rao and Singh studies India. In both cases, it appears that large-scale distortions away from an ideal tax system emerge as a result. Two papers focus on organizational design and tax compliance. The chapter by McLaren examines the idea that corruption can be fought by an increase in functional specialization within a tax bureaucracy. The chapter by Silva studies the design of incentives for the managers of a tax bureaucracy. Finally, The paper by McMillan looks at the problem of opportunistic taxation, with a particular regard to the African context, and studies various ways in which the problem can be alleviated.

A frequent theme in these papers is the inextricable relationship between tax policy and tax administration, emphasizing the point that optimal policy requires simultaneous consideration of the design of the tax code and of the administrative structure created to enforce it.
Acknowledgments

The papers in this volume were presented at a joint Columbia University/World Bank conference held on February 18-19, 2000 at Columbia University, entitled "Institutional Elements of Tax Design and Reform." It was sponsored by Columbia's Program in Economic Policy Management (PEPM). Thanks are due to a great many individuals who made the conference and this volume possible. Michael Engelschalk of the World Bank and his assistant Anna Hansson were energetically involved in putting the conference together. Tuan Minh Le's efforts were crucial in the publication process. Hyejoon Im at the University of Virginia provided excellent editorial help. The support of the Tax Policy and Administration Thematic Group at the World Bank was crucial to the realization of this volume. Luca Barbone of the World Bank and Arindam Das-Gupta, who was at the Bank at the time, provided invaluable guidance to the project in its early stages. The director of PEPM at Columbia University, Francisco Rivera-Batiz, offered enormous support. The PEPM assistant director, Maya Haddow, and her assistant, Deborah Wilson, along with the rest of the PEPM staff, did the hard logistical work that made the event possible. Thanks are due to the PEPM students, as well, for participating and contributing to the discussion. Richard Bird added enormously by his presence and participation, and Dilip Mookherjee also went the extra mile. In addition to the authors listed here, several scholars presented work that for one reason or another is not included in this volume: Keiko Kubota (World Bank), Raymond Fisman (Columbia University), Anwar Shah (World Bank), Dilip Mookherjee (Boston University), Jaime L. Vazquez-Caro (World Bank), John Crotty (IMF), and Jit Gill (World Bank). Several colleagues served adeptly as discussants: Kenneth Leonard (Columbia), Margaret Madajewicz (Columbia), Ronald Miller (Columbia), Douglas Marcouiller (Boston College), Dilip Mookherjee (Boston University), and Rohini Pande (Columbia). Heartfelt thanks go to all.
**Contributors**

Arindam Das-Gupta is at the Indira Ghandi Institute Of Development Research, India.

Ira N. Gang is a member of the economics department of Rutgers University.

Daniel Kobb is the director of K Consult USA in Tanzania.

John McLaren is a member of the department of economics at the University of Virginia.

Margaret S. McMillan is a member of the economics department at Tufts University.

M. Govinda Rao is the director of the Institute for Social and Economic Change, Bangalore, India.

Emilson C. D. Silva is a member of the economics department at Tulane University.

Nirvikar Singh is a member of the economics department of the University of California at Santa Cruz.

Daniel Treisman is a member of the department of political science of the University of California, Los Angeles.
Chapter I
Overview

This is a collection of papers that study the constraints on fiscal systems imposed by problems of institutions, administration, and incentives in developing and post-Socialist economies.

The economics literature in the late 1980's and 1990's saw a surge in analytical work on the design and reform of taxes in developing countries. The collection of papers edited by Newbery and Stern (1987) provided an important impetus and a powerful theoretical toolkit, which defined the agenda for much of what came after. This was a neoclassical approach, which made heavy use of Ramsey rules and of estimated elasticities of substitution for empirical implementation. The framework was used in a variety of fruitful ways. For example, Ahmad and Stern (1991) produced a remarkably detailed study of tax reform in Pakistan, Newbery (1990) analyzed agriculture taxation in Ghana, and chapter 5 of Deaton (1997) analyzes the econometric toolkit, with an application to the tax systems of India and Pakistan.

Despite the power of this approach, however, it omitted a range of important issues by fiat, by effectively assuming that the taxes and transfers under consideration are costlessly enforced. The only exception is the occasional fiscal instrument that is assumed to be administratively so difficult that it is not available at all; for example, it is common in that literature to assume as a crude approximation that income taxes are not an option, citing in particular the difficulty of enforcement in the agricultural sector of a low-income economy. Thus, since the focus of that literature was the tax rates and statutory tax structures, there was a tendency to assume either perfect enforceability or perfect unenforceability, and this squeezes questions of administration out of the discussion by fiat.

However, in the eyes of many observers, reform of tax administration is the first order of business (see, for example, Bird and Casanegra de Jantscher 1992). This has been emphasized in the corruption literature, as exemplified in interdisciplinary work by Klitgaard (1988) and Rose-Ackerman (1999). Theoretical economic literatures on evasion and on corruption have developed that usually (if not quite always) take tax policy as given (see Cowell 1990, Bardhan 1997 and McLaren 1997 for surveys). Das-Gupta and Mookherjee (1998) apply this body of theory to many aspects of tax administration in the Third World. Thus, tax administration has now also received a great deal of scholarly attention. However, it has been approximately true that scholars who were interested in tax administration simply would take tax policy as given, and vice versa.\footnote{There are some exceptions. For two theoretical examples, Hindriks, Keen, and Muthoo (1999) study the effect of corruptible tax inspectors on the shape of the optimal income tax schedule, and McLaren (1998) studies the effect of costly enforcement on the optimal pattern of indirect taxes. There are other examples, but in the research literature this intersection remains underexplored.}

In order to promote a more holistic approach that sees tax policy and tax administration as two sides of a common policy problem within an analytical theoretical and empirical framework, and
in order to promote discussion between economists and non-economists working on these problems, as well as between academics and practitioners, a joint Columbia University/World Bank conference was held on February 18-19, 2000 at Columbia University, entitled “Institutional Elements of Tax Design and Reform.” It was sponsored by Columbia’s Program in Economic Policy Management. The papers in this volume were all presented at that conference. These papers examine institutions and their impact from different perspectives, and can be divided roughly into four categories: Papers on the administration of indirect taxation; papers on politics and incentives in fiscal federalism; papers on organizational design and tax compliance; and a paper on reputation and opportunism in taxation. This essay will provide an overview of the work, taking each of these topics in turn.

Administration of Indirect Taxation

The two chapters under this heading are by Kobb and by Das-Gupta and Gang. Although very different in approach, focus and method, they both emphasize the importance (and complexity) of evasion incentives, and both analyze directions for institutional reform.

The paper by Kobb is a case study of problems of indirect taxation in Tanzania, based on the author’s extensive experience documenting taxation practice as a consultant there and on data he has collected in the field. The paper accomplishes two objectives. One is to show how evasion can be documented and quantified by a variety of clever tricks. One is simple counting: a researcher can in many cases get a good estimate of the size of a tax base, for example, by counting the number of vendors at a public market where vendors are required to pay a fee to the government. This number can then be compared with the number of vendors implied by actual revenue collections. More sophisticated is simply observing tax collection closely for a period, and comparing collections during that period with collections at other times. Based on a presumption that the tax collectors behave with diligence and rectitude while they are watched, the difference (correcting for ordinary fluctuations) can be attributed to laxity or corruption on the part of tax collectors during unobserved periods. This could be called the “Heisenberg” method, since, as Kobb points out, it depends on the modification of the tax collectors’ behavior by the act of being observed. Using these methods on a variety of indirect taxes, Kobb demonstrates a quite gigantic degree of tax evasion, in each case resulting in payment of well under half of the taxes due.

The second objective is to study a particular type of reform aimed at curbing evasion: Franchising, or privatizing the right to tax. This has been tried in several Tanzanian towns in the area of vendor fees for access to a public market, as mentioned above. In these experiments the right to collect fees for a given market was auctioned off to private concessionaires. The winning bidder would sign a contract to pay a monthly fee equal to the amount of that enterprise’s bid. Many unforeseen problems emerged, most notably dramatic problems with non-payment of the rental fee, compounded by sloppy administration that made it difficult to enforce the terms of the contract, and which Kobb speculates might even be deliberate sloppiness induced by a new form of bribery created by the franchising scheme. However, the striking finding is that despite these problems the scheme resulted in more than a doubling of the government’s net revenues from the vendor tax. This is a rare and fascinating case study of a type of policy reform that is routinely discussed in theory but not commonly observed in practice. The paper also discusses the possibility of applying similar techniques to tax collection, and to other areas of governance, in Tanzania more generally.
The chapter by Das-Gupta and Gang is a theoretical study of evasion in a value-added tax (VAT). Under a VAT, a producer owes tax only on its value-added; in practice, this means roughly that tax is paid on all industrial output, including produced inputs, but a purchaser of an input is entitled to claim a refund on taxes paid on the input. VAT’s have propagated widely through world tax systems in recent years, encouraged by international organizations who perceive them as offering both allocative and administrative benefits over alternative forms.

Two purported properties of VAT’s have had paramount importance in policy discussions. The first is that because of the refund they effectively leave purchased inputs untaxed, so that buyer and seller face the same effective price for the input, which is crucial for productive efficiency and hence for optimality of the tax system (Diamond and Mirrlees 1971). The second is that they make it harder to evade tax, because the buyer of an input will have a strong incentive to keep records of its purchase and report it to the government, and therefore the input supplier will not be able to conceal the sale of the input. As a result, it will be difficult for the input seller to avoid tax on its value added.

However, this important pair of ideas has not been subject to careful theoretical analysis, and that is what this paper does. The model takes into account the important fact that the conflict of interest noted above as the second advantage of a VAT depends both upon auditing and upon a further act of cross-checking tax returns of industrial producers with those of their customers, both of which are costly. As a result, these procedures will typically be done only for a sample of returns, resulting in a probability less than one that any act of evasion will be detected. This leaves open the possibility that some evasion will be possible even under a VAT. In addition, if this occurs, so that the second purported property of a VAT noted above fails, then as a result the first property will also be thrown into question, a result that is not at all obvious but potentially quite important. Thus, the two properties widely assumed to hold for VAT’s are not only both open to question, but their potential failures are closely linked.

To see how this can occur, consider the following example. Suppose that the probability that any given input seller will be audited is equal to $q \in (0,1)$, and the probability that any given input buyer is audited is given by $p \in (0,1)$. Suppose that an audit is always completely successful at revealing all of the documents relevant to the firm’s tax liability. (The paper relaxes this assumption.) Suppose that when the auditor finds a seller’s copy of a receipt for sale of an input, it is generally not able to track down the purchaser (again, this is relaxed in the paper), but if it finds the receipt in the hands of the buyer, it will be able to track down the seller with probability $S \in (0,1)$. Assume for the sake of argument that neither firm is reporting any sales to the tax authority, and thus both are fully evading their taxes (in the paper, they may choose any level of evasion). If either firm’s evasions are detected, it will need to pay its full tax obligation, assessed as a constant $(t + f)$ times its value added, where $t$ is the tax rate and $f$ is a fine rate. In this situation, the effective price of the input from the point of view of the buyer is the transaction price minus the expected tax savings resulting from the purchase, which are given by $pt(t + f)$. (The transaction provides no tax savings if the firm successfully evades all taxes, but if it is caught in evasion, which occurs with probability $p$, it receives a benefit from subtracting the input purchase from its value added.) However, the effective price from the point of view of the seller is given by the transaction price...
minus the expected tax it must pay, which is given by \((q + (1 - q)pS)(t + f)\). (Its evasion is discovered either if it is audited, which occurs with probability \(q\), or if it is not audited but its customer is, and the buyer’s receipt is traced to the seller; this occurs with probability \((1 - q)pS\).) Therefore, the effective price of the input differs between buyer and seller unless \(p = (q + (1 - q)pS)\), and there is no reason to expect this equality to hold. (In particular, if \(p = q\) it must fail.) Thus, productivity efficiency is not generally realized when the VAT is not perfectly enforced.

The model is necessarily complicated because the problem is complicated, but the point is very much worth making, because of the widespread influence of the optimistic assumptions made about VAT’s in the policy world. It would be of great interest of the effects in question could be quantified empirically.

**Politics and Incentives in Fiscal Federalism**

Two chapters, one by Rao and Singh and the other by Treisman, examine case studies in fiscal federalism in which the determination of fiscal outcomes is, at least to a considerable and statistically verifiable degree, a matter of bargaining or gamesmanship between political entities in the center and in the periphery. In both cases, it appears that large-scale distortions away from any conception of an ideal tax system emerge as a result.

The article by Rao and Singh focuses on India, in particular on the workings of transfers from the central government to the states. Under the Indian system, this is a crucial form of financing of state expenditure, generally amounting to approximately a third of the total state budget. Rao and Singh examine the transfers from many angles, historical, theoretical and empirical, documenting the complex system under which these transfers are administered.

The framers of the constitution had envisioned a system in which transfers would be administered following recommendations of the Finance Commission (FC), an apolitical advisory body appointed by the central government every five years. However, in practice generally more than a third of total grants have been conferred by agencies outside of the control of the FC, notably by the Planning Commission (PC), which has no expectation of apolitical conduct. Both bodies are constrained by an elaborate set of rules. In the case of the grants conferred by the PC, the rules are ostensibly intended to produce a progressive transfer scheme, with higher per capita grants to poorer states *ceteris paribus*, along with a variety of other criteria, and the FC grants are supposed to bridge the gap between the “needs” of each state and its “available resources.” The authors document the confusions and distortions created by this system of competing federal institutions. However, the striking result comes from the statistical analysis, which shows that the most consistently significant determinants of either FC or PC grants are political. In particular, FC grants per capita to a given state are positively related to the size of the state, whether measured in terms of population or in terms of state GDP. This is most naturally interpreted as a measure of the state’s bargaining power, either in terms of economic clout or of electoral influence. On the other hand, grants approved by the PC per capita to a given state are strongly related to whether or not that state’s government is of the same party as the federal government (or at least a partner in a governing coalition). No other variables are consistently significant. The bottom line is clear: despite the apparent impartiality built
into the system by a complex of rules and institutions, it nonetheless rewards states to a significant degree on the basis of clout.

The chapter by Treisman focuses on Russia. Here the issue is collaboration between provincial governments and businesses to evade taxes owed to the center. Although this is not directly observable, there is abundant indirect evidence of the phenomenon, and the paper builds on earlier work by the author to identify its determinants. Evasion is reflected indirectly by a state’s tax arrears, meaning the portion of taxes owed to the center by enterprises operating in the state that have not been paid (plus fines and interest from previous arrears). These arrears grew rapidly in the 1990s, as the federal government’s share of collected taxes fell. Different provinces’ arrears grew at widely different rates, which allows for empirical analysis to determine what the forces driving arrears were. Despite the extreme data problems (discussed at length in the paper), three determinants emerge as important. First, a state is more likely to have large arrears growth if its firms tend to be large ones. This is interpreted as consistent with the joint evasion story, both because a large enterprise would be more likely to lobby the provincial government effectively for help against central government tax collection, and because the provincial government would be better able to collude with an industry if that required negotiation of a handful of key players rather than with a large number of small ones. Second, a recently elected Communist governor tends to increase a province’s arrears, consistent with the idea that political animosity between the provincial and federal governments would reduce the provincial government’s incentive to help in the enforcement of federal taxes. Finally, and significantly (in the statistical sense as well as the ordinary sense), an increase in the province’s share of shared taxes is associated with a drop in that province’s arrears. Clearly, this is understandable in terms of the collusion story, because a provincial government’s incentive to help an enterprise resist taxes will be curtailed if that government has a large stake in the resulting revenues. This finding provides important evidence for the collusion theory.

The common theme of these two papers is that the institutions of federalism can create a political game between different levels of government (and in the Indian case, between competing institutions of federalism within the central government), and that this game can have an enormous effect on the way the fiscal system actually works.

Organizational Design and Tax Compliance

The chapters by McLaren and Silva fall into this category. The chapter by McLaren examines formally an idea that has appeared less formally in policy discussions: that corruption can be fought by an increase in functional specialization within a tax bureaucracy. If auditors specialize geographically, then each will audit completely a small number of taxpayers. If they specialize functionally, then each will do part of the audit for a large number of taxpayers, working a part of a team. The idea is then that if an audit must be performed by a team of officers, collusion between the taxpayer and the officials will be more difficult to sustain than would be collusion between the taxpayer and a single auditor. A larger conspiracy is more difficult to maintain, at least in part because the benefits must be spread more thinly. The paper presents a simple model in which the tradeoffs are easy to see, and indicates under what conditions such an increase in specialization is likely to be cost-effective.
The chapter by Silva focuses on the other end of the organizational chart. Early on, the economics of tax evasion emphasized the design of incentives for the taxpayer. Later work encompassed incentives for inspectors and auditors as well. It has become recognized, however, that incentives for the management of a tax bureaucracy are also important, as demonstrated, for example, in the dramatic effects of incentive reform in the Brazilian tax system (Kahn, et. al, 2001). The chapter by Silva presents a technical analysis of that incentive problem.

The paper emphasizes the quantity/quality tradeoff for tax enforcement in a context in which managerial time and attention are scarce resources. This is exemplified by the specification for auditing and inspection quality, which is assumed to be an increasing function of the tax authority's effort devoted to quality improvement but decreasing in the quantity of inspections and audits, implicitly because with a larger volume of inspection and audits, managerial time must be spread more thinly. This results in implications similar to the “multi-tasking” literature in incentive theory. For example, if the tax authority is given strong incentives to increase collections from audits, it is possible that this will have an indirect effect of reducing the quality of enforcement, by deflecting managerial efforts away from quality and by raising the volume of audits to which quality management must be deflected.

In addition, the model suggests that in an incentive sense quality and quantity tend to be complements, in the sense that an exogenous increase in quality will, under plausible conditions, raise the number both of audits and of inspections. This suggests that any policy change that induces the tax authority to improve enforcement quality will offer powerful positive effects on enforcement. Such policies include (under reasonable conditions) at least two prominent possibilities. The first is computerization. This reduces the cost of generating a given quality and quantity of enforcement. The second is raising tax rates on upper income brackets. This increases the return the tax authority will enjoy from enforcement for any level of evasion, thus inducing more managerial effort. This implies the perhaps surprising conclusion that increasing high-bracket tax rates in this model will lower evasion; the reason is sharpened enforcement incentives for the tax authority, which are understood by taxpayers and thus act as a deterrent to evasion.

Reputation and Opportunism in Taxation

The paper by Margaret McMillan looks at the problem of opportunistic taxation, with a particular regard to the African context. If investors make a sunk investment, then in a sense they become a captive tax base, allowing the government to tax away their ex post profits and thus preventing them from recouping a positive return. If investors understand this, then they will be reluctant to make the investment in the first place, thus causing a highly inefficient outcome, unattractive for the investors and the government as well. There are various ways around this problem, including reputational solutions and commitment devices. McMillan reviews her own recent research on this, showing that the problem has been a serious one in cash crops in Sub-Saharan Africa, many of which require a substantial sunk investment in planting and nurturing a tree for years before any output is actually realized. Further, the severity of the problem is affected in a variety of predictable ways by changes in the environment such as the importance of the sunk costs in total costs and the future predicted profitability of the crop. This has consequences for domestic
innovative activity and growth rates, as is confirmed in the data. Finally, McMillan extrapolates from the theory and empirical work to analyze the institutional factors that may mitigate the problem of opportunism in taxation in the long run, including political parties, international institutions, and delegation of decision-making.

A frequent theme in these papers is the inextricable relationship between tax policy and tax administration. In the Treisman paper, it is clear that the formulas by which taxes are to be shared between different levels of government have an effect on inter-level cooperation, and hence on compliance rates. In that case, tax policy affects tax administration. Rao and Singh show that the actual fiscal regime implemented is shaped, independently of the statutory rules, by incentives created by the administrative structures in place to implement them. In that case, fiscal administration affects (effective) fiscal policy. Silva shows how the quality of administrative effort in the tax service can be shaped in part by the tax rates it is charged with enforcing. McMillan shows that an African government’s choice of current tax policy can affect its credibility in future, which is a form of institutional capital that can affect its freedom to set tax policy in the future. Such links are unavoidable in the real world, and are promising areas for further exploration.
References


Chapter 2
Corruption in Tanzania: Counting and Franchise Bidding

Daniel Kobb

Corruption in Tanzania is endemic and has invaded virtually all aspects of daily life. Recognizing the severity of this problem, the President, on January 17, 1996 appointed a special "Commission on the Causes of Corruption in the Country." The Commission primarily investigated the giving and receiving of bribes and favors (but not outright theft, forgery, and other forms of white-collar crime, which are also rampant). The final document, referred to as the "Warioba Report" (after the Commission's chairperson, Joseph Warioba, a former Prime Minister) documents 127 general "types," "cases," "areas," or "forms" of corruption, typically at the level of abstraction of: "policemen solicit bribes from offenders to let them free" (page 49). Other government evaluations are similarly unambiguous. For example, the Presidential Commission of Inquiry into Land states that "evidence also points to sheer lack of honesty and integrity on part of land officers in the administration of land...the commission is convinced that there is a serious erosion of ethics, professional discipline, and elementary integrity" (p. 100). Using a "poll of polls" and based on interviews with businessmen and the general public, Transparency International, in 1998, rated Tanzania as the world's fourth most corrupt nation in its set of 85 countries; in 1999 it was the seventh most corrupt nation out of a set of 99 countries.1

Without the aid of movie cameras and tape-recorders or without confiscating forged documents corruption is difficult to prove. In this paper I focus on cases where actual tax collection differs so radically from potential tax collection that corruption should be implied.2 Other sources of the gap between actual and potential revenue include a lack of effort, inability or ineptitude, the high cost of collecting taxes (where costs are both political and economic), insufficient inputs, and difficulties in monitoring and supervision (and therefore a lack of accountability). In laymen's terms, taxes may be under-collected both intentionally (through corruption) or unintentionally (through inability, and a host of other factors). Intent may matter legally, but from the perspective of revenue generation the distinction may be less relevant. In fact, since inability (poor performance and incompetence) does not generally constitute serious grounds for termination in Tanzania, corrupt officials are motivated to create disorganization and

---

1 I have benefited from discussions with Stig Johansson, Hans Lipp, John McLaren, and Dani Rodrik. Bert Koppen helped with earlier drafts and a related paper. All errors and views expressed remain the responsibility of the author.
2 In 1998 Tanzania was tied for fourth worst with Nigeria. In both 1998 and 1999 Tanzania had a score of 1.9 (with ten as clean, zero as dirty). The highest score, 10, was received by Denmark in both years. Tanzania was not in the survey for 1995-7.

2 In the cases I discuss, interviewees hinted that corruption was a problem, and there were other forms of evidence that corruption might be at hand.
intransparency. Some deliberately under-perform in order to cloak corruption as financial uncertainty, poor planning, unrealistic budgeting, inadequate record keeping, irregular accounting, overlapping responsibilities, and institutional fragmentation. In this sense poor tax collection is symptomatic of a much larger problem of weak governance, where neither term in the phrase "public service" may longer apply. Klitgaard (1995) describes organizations with these qualities as "sick institutions."3

This paper discusses some recent field experiences in Tanzania and tells a rather simple story. Donors in Tanzania have been increasingly concerned about the low financial contribution from their host-government counterparts. I have undertaken several consultancies (financed by donors but on behalf of local governments and Ministries, such as the Ministry of Natural Resources and Tourism) intended to increase tax revenue. During the course of this work it became apparent that most taxes in Tanzania were avoided, that the problem of tax collection was one of poor enforcement not inadequate legislation. Though the estimates I present are imprecise, quantification has proven to be an important first step in moving policy makers away from perception (taxes are evaded) toward concrete reform. My second goal was to convince clients that "poverty," a lack of funds, and "insufficient capacity" (for example, a lack of equipment) were excuses rather than serious constraints to tax collection. In the cases I document, it does not take money to collect money. As an experiment, and on a small scale, clients were encouraged to create tax "franchises" or "concessions," whereby private sector concerns were auctioned the right to collect taxes on behalf of the government. My results are preliminary and caution is obviously warranted. Nonetheless, I found that privatizing tax collection increased net revenues by 129 percent, that this is consistent with a focus on improved enforcement and incentives, and that there are viable alternatives to the "capacity building" input-oriented approach to improving governance.

In the remainder of this paper I do not attempt to explain these results in a full and coherent model of "insubordination" whereby superiors attempt to regain control over tax streams. Such a model would need to be quite detailed in order to preserve internal consistency. With this limitation in mind, I proceed inductively. Section one spells-out a simple methodology for implying corruption. Section two provides two examples of "corruption detection" and their implications. Section three identifies some of the theoretic concepts I found helpful in explaining franchise bidding or privatization. In section four I summarize, extend, and speculate.

I. A Methodology: Simple Counting

In the case of unit or specific taxes it takes little more than counting persons, commodities, or events to imply corruption. In general, this methodology reconciles paper and reality by comparing historical records (actual tax collection) with observations during a sample

3 "Sick institutions are those in which a substantial number of employees do not come to work or do other work (or nothing at all) while there; where corruption and favoritism are not isolated instances but the corrosive norm; where pay scales in real terms have collapsed so that low and middle level employees cannot feed and house their families on official pay; and where employees therefore seek other forms of compensation, including travel, study allowances, non-wage benefits as well as illicit payments for doing (or not doing) their public duties. Sick institutions do not function...Public service becomes a source of public embarrassment or indignation....Good performance goes relatively unrecorded and bad performance is relatively unpunished. Information and evaluation are scarce and expensive...Some governments face soft budget constraints, meaning that foreign aid will fill a good part of any deficiency resulting from inefficiency or corruption" (selected pages).
I focus on the incongruity of particular observable transactions. One popular form of cheating (which I do not investigate) involves over-counting. In this case, personnel officers might employ non-existent staff ("ghost" workers) and pocket their salaries or officials might be paid for taking fake ("ghost") trips or may purchase non-existent goods ("air").4 In this paper I focus on undercounting, where what is recorded on paper is a fraction of what exists in reality.

Two examples will better elucidate this. In Tanzania sellers at markets are assessed a fixed daily fee (ushuru in Swahili). The typical rate is 100 Shillings per person per day (there are now roughly 800 Shillings to the US dollars). According to its audited 1995 accounts, the Korogwe District Council collected on average roughly 260,000 Shillings per month from the Manundu Market. There was one problem: revenues implied there were on average 86 sellers physically present at the market each day (paper), while visual observation (counting) confirmed there were closer to 280 sellers per day (this I consider the "reality").5 In revenue terms, potential gross revenue was 860,000 Shillings per month, not 260,000 Shillings per month. Roadblocks provide a second way to "count" and sample. In Tabora Region (Central Tanzania), enumerators spent 19 days simply watching and recording traffic flowing into the town center (Koppers 1997). They simply counted open lorries of timber and charcoal. In all, they observed 109 lorries. Using standard capacities (for example a 7-ton lorry can carry about 140 bags of charcoal), the average daily revenue during the sample period (based on traffic and tax rates) was calculated to be 970,000 Shillings. Unfortunately, the average daily revenue according to collection statistics was only 96,000 Shillings.

Counting should be considered a preliminary step only. Where bribery has permeated the tax collection system, the best way to estimate revenue potential is to temporarily substitute the current tax collection regime with an outside team of honest and motivated individuals; that is to "collect yourself" for a sample period and then to compare "your" results with previous figures. On a limited basis, the Mwanga District Council employed this method. For example, and on average, daily taxes on the transportation of sand in 1997 were 10,315 Shillings; when a special revenue team took over for one day, revenue "jumped" to 70,500 Shillings (Mwanga District Council, 1998).6 A related, but slightly more sophisticated method is to transform the usual social science problem of the interfering observer into a natural experiment. During a sample period accompany tax collectors on their daily rounds and then compare collection during and in the absence of external observation.7 This second method ensures that the gap is due to corruption or a lack of effort, not to alternative techniques or to varying levels of inputs. For example, when I (Kobb, 1997a) observed revenue collection at the bi-monthly market in Msingisi Village (Kilosa District) revenues were 72,800 Shillings; in my absence, two weeks

4 See Ruzindana (1997) for a brief description of the problem in neighbouring Uganda.
5 It would be stretching the imagination to say that tax collectors were unable to see the sellers or that they required better inputs or work conditions to collect the tax.
6 This is based on only observation, but the amount of revenue lost is substantial. If the figures above are indicative, losses on the sand tax are equivalent to 20% of total revenue. Second, as with other checkpoints, there is a good deal of seasonality. Finally, transporters will undoubtedly change their behaviour should a more efficient tax collection body be in place.
7 The observer should be a person of "standing" (i.e. should somehow "intimidate" the collector) but should not be directly involved in the collection process. The idea is to provide the requisite motivation. I (Kobb, 1997a) employed the following steps. I introduced myself to the Revenue Collectors as an employee of a particular Embassy, assigned to research tax collection. I noted the collectors' names (dutifully entered into a notebook) and looked through receipt books, carefully recording all starting and finishing numbers. I sympathised with the difficult job revenue collectors had, but did not speak to taxpayers nor did I advise tax collectors. While accompanying Revenue Collectors on their rounds I took fictitious notes.
earlier, 33,800 Shillings was collected. Sustained over one year, simply "observing" this one market (of which there are many in the District) would increase annual revenues by 2 percent.

Academic economists may object to this methodology on the grounds that it only generates estimates. On the other hand, like all sampling, it is amenable to statistical analysis, for example, in generating confidence intervals. Despite problems in reproducibility (stemming from sampling, human error, seasonality, and the need to make a few assumptions along the way), I believe such figures more closely describe reality than do official statistics (which are also estimates and are also derived from interviews, for example between farmers and agriculture extensionists or taxpayers and tax collectors).

II. Two Cases and Some Limitations

Markets

Tax franchises provide a possible solution to corruption. To recap, actual tax collection figures at markets in Tanzania conflict with the number of sellers physically present. In response to these problems Korogwe District Council auctioned the right to collect revenue from its main markets. It created a tax concession or "privatized" or "tendered" the market. The right was originally sold for 500,000 Shillings per month whereas net tax collection previously stood at roughly 160,000 Shillings per month. In all I have partial data for 37 markets in 5 districts. Variables include: the actual gross tax collection during a baseline period (28 markets, 57 months) while the market was in public hands, the cost of tax collection (10 markets) while the market was in public hands, potential revenue (derived by counting the number of sellers present; 19 markets), payments due once markets were privatized (the contractual bid, 29 markets, 386 months), and the actual amounts paid by the private concessionaire (26 markets, 383 months). Wherever possible, actual (historical) gross collection figures were calculated by taking an average for the three-month period preceding tax franchising. This data was provided by the treasurer of each District, either from the cashbook (larger markets) or from original receipts. In the case of smaller, more distant markets revenue collectors also collect local liquor licenses, livestock fees and other taxes. Payments are presented by hand to the accounts department, often as a single transaction and containing breakdowns scribbled on the back of a receipt. The cost of tax collection in Tanzania is very high; since most institutions in Tanzania do not use cost center accounting, the best one can do is estimate. For the 10 markets in which there is data, I went through a sample week or month with District Treasurers to track all direct expenditures. The largest component was the salary of the collector, but in some cases

---

8 This type of data is amenable to all sorts of statistical analysis for testing means and confidence intervals. In the case of this market, once the exercise was completed I returned to the market and randomly checked 20 businesses for receipts. Of these seventeen could provide me with copies of their receipts. The market is highly unorganized and difficult to patrol systematically. Many sellers hide as soon as they see revenue collectors entering their area, frequently leaving young children to tend their businesses. Not a single customer paid willingly without complaining and much argumentation. Force was used routinely and collectors had to confiscate goods in the hope that businesspersons would follow them up and pay. Others claimed they "had no money" and collectors had to search their goods for the hidden cash (usually finding it amongst their goods). Lying was prevalent; it was the norm not an exception. Rates were routinely negotiated, making the payment of tax entirely arbitrary. In other cases, neighboring sellers attempted to make joint fee payments, in essence pooling their taxes and sharing a receipt. Many sellers therefore claimed to be in partnerships with friends or relatives or to be selling on behalf of richer owners. Others sold quickly and left the market before revenue collectors could assess them.
(Kwamshemshi market) the exorbitant transportation costs involved were derived using a flat cost per kilometre.

In all, 29 markets in three districts have been privatized. This took quite some time and required lengthy discussions during quarterly planning and finance committee meetings. Recommendations for privatization had been made by September 1996 in a total of five districts. Three councils took up the idea: Korogwe in August 1997, Lushoto in January 1998, and Mwanga in August 1998. Korogwe and Lushoto were revisited in January 1999 to update the information provided in this paper.

Table 2-1: Ratios of Revenues before/after Tax Concession

<table>
<thead>
<tr>
<th>Numerator</th>
<th>Denominator</th>
<th>Markets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross Public Net</td>
<td>(n)</td>
</tr>
<tr>
<td></td>
<td>Collection</td>
<td>Collection</td>
</tr>
<tr>
<td>Potential Revenue</td>
<td>4.48</td>
<td>7.47</td>
</tr>
<tr>
<td>Kilosa</td>
<td>2.12</td>
<td>3.73</td>
</tr>
<tr>
<td>Korogwe</td>
<td>5.31</td>
<td>8.37</td>
</tr>
<tr>
<td>Lushoto</td>
<td>5.06</td>
<td>8.71</td>
</tr>
<tr>
<td>Mwanga</td>
<td>3.73</td>
<td>6.22</td>
</tr>
<tr>
<td>Ulanga</td>
<td>4.55</td>
<td>7.99</td>
</tr>
<tr>
<td>Accepted Bid</td>
<td>2.11</td>
<td>3.54</td>
</tr>
<tr>
<td>Korogwe</td>
<td>2.49</td>
<td>3.98</td>
</tr>
<tr>
<td>Lushoto</td>
<td>2.13</td>
<td>3.70</td>
</tr>
<tr>
<td>Mwanga</td>
<td>1.50</td>
<td>2.47</td>
</tr>
<tr>
<td>Amount Paid</td>
<td>1.36</td>
<td>2.29</td>
</tr>
<tr>
<td>Korogwe</td>
<td>1.94</td>
<td>3.10</td>
</tr>
<tr>
<td>Lushoto</td>
<td>1.03</td>
<td>1.80</td>
</tr>
</tbody>
</table>

Source: Data gathered by author.

To evaluate the success of tax concessions, I construct simple ratios using the above variables. To start, the tax collector is assumed to be interested in total tax collection, rather than representative markets: in other words, the successful privatization of a large market can more than compensate for numerous minor failures. Table 2-1 shows that for the 19 markets with data on potential revenue (counting) and baseline public collection figures, gross public collection was only 22 percent of the potential revenue. That is, the ratio of total potential revenue to total baseline gross public revenue is 4.48. When markets were privatized, the contractual bid was 2.11 times baseline gross collection. Of these 20 markets, there is only one case in which the contractual bid was less than its public gross baseline revenue. As a final note, this may understate the potential of franchising: in one district (Mwanga), the highest bid was not routinely accepted because the council had a high outstanding debt with the bidder.

---

9 To provide some rough background, there are roughly 120 districts in Tanzania. A typical rural district (local government) will be governed by about 30 elected ward councillors. Councillors undertake all legislative activities. The Central Government, through the Civil Service commission and relevant ministries appoints roughly 100 civil servants to execute a district’s policies.

10 Information on bids was contained in Planning and Finance Committee Minutes. In Mwanga, the council did not want the highest bidder to deduct the debt repayment directly from his monthly rental fees.
The relevant comparison is however between actual payments under concessionaires and actual net payments under public ownership. Market fees are extremely expensive to collect, and I have cost data on only 10 markets under public control.\(^1\) Using this cross-section, I regressed cost on revenue, trying four functional forms. For each market, I then calculated the minimum cost using these four alternatives.\(^2\) Again referring to Table 2-1, the contractually accepted bid was 3.54 times net public collection.

![Figure 2-1: Payment History](image)

While successful, much of the potential of tax concessions was not realized. Many concessionaires abrogated their contracts; councils issuing the concession have failed to take decisive action. In rough terms, only 60 percent of what is due has been paid (see Figure 2-1).\(^3\) Several markets are problematic, but the constraint is clearly willingness not ability to pay. Concessionaires at these markets continue to collect revenue from the public but steadfastly refuse to remit these funds to the government. Typical patterns are exemplified by the markets at Kwemakame, Baga, and Lukozi (in Lushoto District). As can be seen in Table 2-2, payments dropped to zero from July to December 1999.

Interviewees at Lushoto District Council claim legal proceedings have been initiated. Although Korogwe seems to have selected more reliable concessionaires (80 percent of its contractual payments have been made, compared with 42 percent for Lushoto), administrative problems have nevertheless emerged. According to one interviewee in the accounts department, councilors regularly stop by one market to raid its coffers. In 1998 the council could not sue for arrears in several markets since “the legally binding agreements have been lost.”

Despite payment problems, there is a strong indication that net revenues have increased: looking at Table 2-1, privatized markets bring in 2.29 times the net revenues public markets did.

\(^{11}\) An advantage of tax franchises is improved data collection: councils could easily provide data on monthly revenues for markets once they were privatized; while publicly run, costs and revenues were available only through primary documents and much calculation.

\(^{12}\) The simple functional forms, using monthly collection costs (C) and monthly revenues (REV) were: linear with intercept (C=36727+0.20*REV; \(R^2 = 0.336\)), linear with no intercept (C=0.43*REV; \(R^2 = -0.21\)), Logarithmic (C=-175400+20471*ln(REV); \(R^2 = 0.332\)), and power (C=44029+5E-7*REV\(^2\)+0.686*REV \(R^2 = 0.355\)). All calculations were done in Excel.

\(^{13}\) Councils are admittedly in a learning process and contracts have been improved. Contracts however do not require the signatory to provide collateral, though most probably have the means to (typically an automobile). As can be seen from the graph, the councils themselves have been expanding the experiment, an indication that they find it beneficial.
Much of these gains occurred because administrative costs have in theory been passed onto concessionaires. However, one cannot be completely sure; the administrative costs of markets were primarily salaries, and while some employees have been made redundant, others remain employed in other capacities.

<table>
<thead>
<tr>
<th>Item</th>
<th>Baga</th>
<th>Lukozi</th>
<th>Kwamekame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount Due</td>
<td>150,000</td>
<td>180,000</td>
<td>415,000</td>
</tr>
<tr>
<td>Amount Paid</td>
<td>150,000</td>
<td>150,000</td>
<td>---</td>
</tr>
<tr>
<td>Feb-99</td>
<td>150,000</td>
<td>150,000</td>
<td>415,000</td>
</tr>
<tr>
<td>Mar-99</td>
<td>150,000</td>
<td>50,000</td>
<td>415,000</td>
</tr>
<tr>
<td>Apr-99</td>
<td>0</td>
<td>10,000</td>
<td>300,000</td>
</tr>
<tr>
<td>May-99</td>
<td>0</td>
<td>90,000</td>
<td>215,000</td>
</tr>
<tr>
<td>Jun-99</td>
<td>0</td>
<td>0</td>
<td>240,000</td>
</tr>
<tr>
<td>Jul-Dec 99</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: All figures are monthly, in Tanzanian Shillings.

Privatization has also lead to innovation, as was predicted by Shleifer (1998). In Mwanga Town, one of the first actions of the concessionaire was to erect a gate around the market’s entrance (in decades of management the government failed to do this). Similarly, Korogwe now collects from seven small markets it previously ignored due to high administrative costs. In other instances, the change of governance structure is threatened by taxation. When tax collection is public, the service is untaxed; when privatized, concessionaires are subject to a 20 percent Value Added Tax, an arrangement which may make privatization unviable. In some areas, concessionaires have been given government receipts for issuance, to dissuade the tax authorities from assessing them.

Forest Products in Tanga Region

In Tanzania, forest products (charcoal, timber, poles) are levied a specific tax rate (largely depending on the type of wood). Though the tax is supposed to be paid at the point of production (cutting or burning) evasion is widespread. Compliance rates are often below 5 percent (Koppers, 1997, Silviculture 1991). The problem with charcoal (which I describe below) is probably more typical than that of markets; its smuggling from rural production points to urban consumption points is lucrative, highly politicized, and dominated by the army, police officers, and (frequently) local government officials.

Back of the envelope calculations can be used to indicate the severity of the problem. A simple manipulation of official forestry statistics from Tanga Region shows that the average

---

14 These markets have not been included in the above analysis, although they comprise 7 percent of the average monthly payments made in the database. The idea of this paper is to see whether changes in regimes, from public to private increases tax collection.

15 At this point a reference to low salaries needs to be made. Even a fairly high ranking civil servant in Tanzania will earn a salary between $100-150 per month (very roughly). Many are posted in somewhat remote areas, where there are limited opportunities for businesses outside the office. In many cases, in fact, the only opportunity lies in the natural resource sector (and herein, their paths cross with the environmentally sensitive donor).
sawmill processes only 3.8 planks of hardwood per day. Paper and reality also conflict along the main trunk road leading to the capital; I counted 985 bags of charcoal openly for sale along the roadside though average taxes paid are equivalent to 33 bags sold per day.\textsuperscript{16} These figures led to a more expensive (though still limited) study in which three roadblocks (manned by police) and one checkpoint (with no police support) were established for a period of three days (all this is originally documented in Kobb, 1998b). The complications then began: since the study was financed by a donor organization integrated into the regional catchment forestry department, foresters were used to undertake the research. Not surprisingly they proceeded to undermine it, by informing some timber dealers that an outside team was investigating the market.\textsuperscript{17} When I attempted to reconcile sampling data from the roadblocks and checkpoints with actual tax collection, it became obvious that primary financial documents (receipts) conflicted with summaries, such as monthly tax collection reports submitted to the Ministry.\textsuperscript{18} Graphs of daily revenues revealed gaps and anomalies, as did a listing of receipt book series in use. A quick Ministerial audit revealed that at least one tax collector had been issued numerous receipt books but could not account for them; he managed to flee with what has been estimated to be almost half the Region’s income due from charcoal.\textsuperscript{19}

Annualized data from the roadblocks indicated potential charcoal royalties to be on the order of 177 Million Shillings per year, while actual payments were 31 Million Shillings per year (see Table 2-3). It is informative to account for this gap of 146 Million Shillings. First, the act of observation more than “doubled” annual revenues from 31 Million to 83 Million Shillings. That is, on an annualized basis, revenues jumped during the observation period due to an increase in voluntary compliance and remittance. The Amboni checkpoint is indicative; we observed this checkpoint for four non-contiguous days during the month of September 1998; miraculously, the four highest collection days of that month corresponded to our days of observation (my calculations of this being a random event are over 1 in 30,000). Army and private commercial vehicles which refused to stop account for the remainder of the unpaid royalty.\textsuperscript{20}

\textsuperscript{16} Depending on marketing strategy, roadside charcoal, at any point in time, may include inventory, or merely be a sample of stocks kept in homes, the forest, or the kiln. In some cases, buyers purchase from roadside wholesalers, in others they place orders and purchase out of site charcoal. For what its worth, a regression of amount sold (according to interviews) and amount observed (as the independent variable) with 9 observations and forced through a zero intercept has the functional form \( y = 1.505x \) with R-Squared 0.7817, implying significant stocks out-of-site.
\textsuperscript{17} In fact, if one wants to imply corruption, it can be done either by calculating revenue gaps, or by trapping civil servants in collusive behavior. Collusion can be of two forms; either businesspersons are told to “reduce” their businesses or businesspersons are told to ensure that all transactions are carried out legally during the observation period. Having involved the forestry department in the planning of the exercise, we then “neglected” to inform them that we would also observe (spy on) three random sawmills for a period of three days. These sawmills were subsequently visited by one district and one regional forester. After the regional forester spotted our enumerators, no more deliveries of wood products were made.
\textsuperscript{18} We employed enumerators (high school students) to do daily sums of receipts (in each day 50-100 receipts were issued; this was summed over the year).
\textsuperscript{19} Exact figures in the report are confidential. But, this raises the issue of why no one noticed a sum as large as this.
\textsuperscript{20} Combining “counting” with interview and other data, tax compliance on all forestry products in Tanga Region is estimated to be 13% (Kobb, 1998b). Again, since transporters may be aware they are being observed during a sample period they may reduce their activities; compliance levels may therefore be on the high side.
Table 2-3: Annual Charcoal Royalties
(Millions of Shillings)

<table>
<thead>
<tr>
<th>Checkpoint</th>
<th>OBSERVATION PERIOD</th>
<th>ACTUAL DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amount Due</td>
<td>Amount Paid</td>
</tr>
<tr>
<td>Amboni</td>
<td>119</td>
<td>71</td>
</tr>
<tr>
<td>Mizani</td>
<td>33</td>
<td>12</td>
</tr>
<tr>
<td>Manga</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>177</td>
<td>83</td>
</tr>
</tbody>
</table>

Notes: All values in Millions of Tanzanian Shillings. Compliance is therefore estimated to be 18 percent (31 ÷ 177). "Actual Data" is the sum of all receipts in the possession of the Regional Catchment Forestry Office from 1 September 1997 to 31 August 1998. “Amount Paid” is the annualized sum of all receipts observed during the monitoring of 4 checkpoints (i.e. scaling sample averages to yearly figures and accounting for seasonality. “Amount Due” is the annualized quantities of charcoal counted by enumerators, which includes vehicles with visible loads that did not stop.

Again, the magnitude of these “inconsistencies” is staggering. If observed charcoal consumption levels in Tanga prevailed in the more populous capital, Dar es Salaam, royalty in Dar would be TSh 1.27 billion per year. This figure is higher than total royalty collected on all forest products nation-wide (TSh 1.1 billion from July 1997 to May 1998, Forestry and Beekeeping Division, 1998, page 3).

The “Tanga Report” took over one year before it was openly addressed at a meeting between the Forestry and Beekeeping Division (FBD) and the donor community. A second revenue study has since been commissioned to investigate the possibility of privatization. Such a concession is bound to be problematic since the trade is a rough one and sufficient police powers will need to be granted to a private company.

III. Franchise Bidding, Insubordination and Rent Seeking

Transactions costs (Coase, 1937) determine whether a firm should produce internally (by authorizing a department to do so) or to purchase externally (by contracting in the market). In a similar vein, governments must also decide whether to produce internally (in this case to collect taxes themselves) or to sub-contract with the market. In the case of markets outlined above, transaction costs arose in the principal-agent context: it is very expensive to monitor and to discipline corrupt subordinates, and superiors themselves may not be motivated to do so. It is also very expensive to enforce contracts in Tanzania, since the legal system is inconsistent and subject to procedural stalling. A consistent model of transactions costs, tax franchises, and insubordination (whereby lower level employees can flagrantly disregard their superiors) needs to contrast the incentive structure between the public and private sector and to fully describe the many options a superior may have for rent seeking, before concluding that he or she will agree to
a policy of tax concessions. In simple English, when the system itself is corrupt, can you really beat it?

In Tanzania, principal-agent corruption also afflicts the private sector, where (as in the case of tax franchising) it is occasionally combated by making the concessionaire the residual claimant. To provide a single example, bus owners frequently sub-contract or rent their vehicles to drivers and turn-boys who remit the owner a fixed daily amount. In this case, daily profits cannot be monitored by the owner (unless he or she accompanies the bus on all its trips) since expenses are easily over-invoiced (fake receipts are common), “air” is pumped into the petrol tank, luggage and goods are transported for unknown fees, and repairs and bribes are made along the way, each in the absence of reliable documentation. Through experience and through trial and error owners know roughly what their incomes should be, but actual income, on any given day, is subject to a good deal of noise.

In a world of tax franchises, the superior has at least four possible sources of additional income (over and above his or her salary). First, the subordinate can continue to collect tax revenues and the two can have an imperfect and informal contract describing how the bribery proceeds will be split. Second, the superior could encourage a competition amongst subordinates for the right to collect taxes and extort bribes from taxpayers. In this case subordinates bid for lucrative postings (this, for example, reputedly occurs with customs jobs). Third, external auctions can be rigged: on the basis of who pays the highest bribe, the superior can award a tax collection franchise to a private external firm. Fourth, and this seems to be the case in my research, superiors can increase their institution’s revenues and then proceed to legally skim them off, in the form of wasteful expenditures. Using this technique, public assets are informally converted to private personal benefits by holding meetings, awarding training, and undertaking unnecessary travel (the sick institution problem which raised the possibility of auctions in the first place). To state the obvious, in the absence of expenditure reform, tax concessions may be welfare reducing.

Though local government accounts in Tanzania are not fully standardized, expenditures across jurisdictions are mainly limited to the payment of salaries for administrators (“line item” 1101 in the Tanzanian accounting system), office expenses, salaries of elected councilors (1554), auto repairs (1402), and travel expenses (1202, referred to as “allowances”). Development activities (capital expenditures) are grouped into certain line items and are given a code above 2800. Average expenditure patterns across 10 rural district councils are summarized in Table 2-4, below.
### Table 2-4: Expenditure Classification for 10 District Councils

<table>
<thead>
<tr>
<th>Expenditure Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Mean Admin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salaries</td>
<td>48%</td>
<td>40%</td>
<td>64%</td>
<td>52%</td>
<td>62%</td>
<td>69%</td>
<td>41%</td>
<td>49%</td>
<td>36%</td>
<td>50%</td>
<td>35%</td>
</tr>
<tr>
<td>Administrators</td>
<td>30%</td>
<td>26%</td>
<td>47%</td>
<td>36%</td>
<td>51%</td>
<td>33%</td>
<td>33%</td>
<td>24%</td>
<td>24%</td>
<td>32%</td>
<td>4%</td>
</tr>
<tr>
<td>Others</td>
<td>6%</td>
<td>6%</td>
<td>2%</td>
<td>2%</td>
<td>3%</td>
<td>0%</td>
<td>3%</td>
<td>7%</td>
<td>9%</td>
<td>9%</td>
<td>4%</td>
</tr>
<tr>
<td>Councilors</td>
<td>11%</td>
<td>8%</td>
<td>15%</td>
<td>14%</td>
<td>8%</td>
<td>8%</td>
<td>13%</td>
<td>5%</td>
<td>9%</td>
<td>9%</td>
<td>10%</td>
</tr>
<tr>
<td>Travel</td>
<td>16%</td>
<td>13%</td>
<td>8%</td>
<td>18%</td>
<td>13%</td>
<td>11%</td>
<td>12%</td>
<td>25%</td>
<td>20%</td>
<td>22%</td>
<td>16%</td>
</tr>
<tr>
<td>Running Costs</td>
<td>22%</td>
<td>22%</td>
<td>15%</td>
<td>20%</td>
<td>13%</td>
<td>13%</td>
<td>29%</td>
<td>15%</td>
<td>28%</td>
<td>16%</td>
<td>3%</td>
</tr>
<tr>
<td>Tax Collection</td>
<td>6%</td>
<td>4%</td>
<td>1%</td>
<td>2%</td>
<td>0%</td>
<td>4%</td>
<td>5%</td>
<td></td>
<td></td>
<td></td>
<td>11%</td>
</tr>
<tr>
<td>Office Expenses</td>
<td>13%</td>
<td>10%</td>
<td>10%</td>
<td>16%</td>
<td>5%</td>
<td>4%</td>
<td>18%</td>
<td></td>
<td></td>
<td></td>
<td>11%</td>
</tr>
<tr>
<td>Others</td>
<td>3%</td>
<td>8%</td>
<td>4%</td>
<td>2%</td>
<td>8%</td>
<td>5%</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td>5%</td>
</tr>
<tr>
<td>Payment of Debts</td>
<td>0%</td>
<td>8%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
</tr>
<tr>
<td>Auto Repairs</td>
<td>11%</td>
<td>9%</td>
<td>8%</td>
<td>10%</td>
<td>7%</td>
<td>6%</td>
<td>11%</td>
<td>10%</td>
<td>11%</td>
<td>13%</td>
<td>10%</td>
</tr>
<tr>
<td>Clearly</td>
<td>3%</td>
<td>8%</td>
<td>5%</td>
<td>0%</td>
<td>5%</td>
<td>1%</td>
<td>7%</td>
<td>1%</td>
<td>4%</td>
<td>0%</td>
<td>3%</td>
</tr>
<tr>
<td>Development</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>67%</td>
</tr>
</tbody>
</table>

Notes: 1 = Lushoto District Council (1998); 2 = Handeni District Council (1998), 3 = Mwanga District Council (1998); 4 = Korogwe District Council (1997); 5 = Muheza District Council (1997), 6 = Kilosa District Council (1997), 7 = Ulanga District Council (1996), 8 = Morogoro Rural District Council (1996), 9 = Kilombero District Council (1996), 10 = Pangani District Council (1995). "Admin" = expenditures that are clearly administration; "?" indicates expenditures that may be either development or administration (and are typically a mix of the two).

### IV. Summary and Final Comments

In Tanzania what exists on paper is often not an accurate reflection of reality. As I have hinted, corruption flourishes in an atmosphere of intentional misinformation, where it can freely masquerade as low motivation and a lack of inputs. To review what transpired, in my sample of markets the overall compliance rate was 22 percent. Magnitudes are important since they essentially rule out the possibility of incentives in the form of commissions, whereby the collector retains something like 5 percent. Where the risk of firing is negligible this type of incentive scheme would have to be so lucrative as to essentially legalize the existing system of corruption; that is, rents tend to dissipate. The general solution to controlling subordinates is to first develop approximate rules of thumb (rough performance indicators) and then to enforce these indicators by shifting risk to the residual claimant (write a contract based on these indicators). I suspect private businesses in Tanzania rely more heavily on targets, sub-contracting, and leasing than do businesses in most other countries.

The Government of Tanzania is clearly moving in the direction of targeting. For example, there are plans to convert some government departments or divisions into Executive Agencies. These will operate outside the civil service system. Taxes are collected by the semi-autonomous Tanzania Revenue Authority (TRA), whose management has their salaries linked to annual collection figures. Finally, a foreign specialist firm carries out pre-shipment customs
To be consistent, one needs to ask: if poor revenue collection can be addressed by imposing hard collection and expenditure targets (in the form of legally binding and enforceable contracts) in an environment of corruption, where do “hard targets” come from?22 Hopefully from a few honest civil servants: an advantage of franchise bidding is that these few officials may have the power to overcome layer upon layer of bureaucratic opportunism. A government relying on franchise bidding for both revenue collection and the supply of public services can indeed be quite small.

In this paper I have tried to move away from a strict definition of corruption: when only 18 percent or 22 percent of taxes are collected (my two examples) is it poor compliance or corruption? The taxpayer has two recourses to a tax he or she doesn’t wish to pay; first avoid the tax, second, if caught, pursue a negotiated settlement. Both of these are practiced widely in Tanzania. From the perspective of the tax collector, the objective is to “pin” the taxpayer, to amass evidence showing the taxpayer to be in violation of the law. But laws are fluid. Where corruption arises due to the principal-agent problem, a simple solution is to simply redefine the transaction: once the concessionaire pays the government for the right of tax collection, he or she is free to negotiate at will with the taxpayer. By definition, whatever fraudulent act occurs at market level, between a private taxpayer and private tax collector is no longer classified as corruption. One way to reduce crime is of course to decriminalize certain activities.

The case of privatized markets revealed some noteworthy divergences between economic theory and economic practice. One would expect competition amongst bidders to reveal information on the market’s value; although the number of sellers is readily observable, no bidder bothered to calculate potential profits based on even a rough census. When interviewed after winning their bids, businessmen mentioned only a general intuition that managing markets was a lucrative affair. Some revised this intuition after a few months of actual operation and successfully petitioned the auctioneer to lower their bids. Secondly, the auctioning body drew up very incomplete contracts with bidders; none demanded security nor specified the consequences of non-payment (such as reclaiming control of the market). Third, the private sector in Tanzania differs significantly from the type of firms portrayed in most microeconomics textbooks; firms in Tanzania are frequently shortsighted and fraudulent. This problem has haunted privatization efforts at the national level and has contributed toward poor loan repayment rates. Fourth, if the problem of corruption is really one of inadequate salaries, this is not borne out by the data. Concessionaires pay their employees less than the public sector, but still apparently out-perform them in terms of revenue collection. The private sector monitors more closely. Perhaps more importantly they employed young primary school leavers (with low expectations) while the public sector employs well-educated, middle aged workers (with high expectations and bad habits). Finally, in the end, many of the contractual gains made under the tax concession were eroded by real or imagined payment problems. These problems may themselves involve corruption, for example, payments to lose documents or agreements not to take action. When a

---

21 This covers all shipments valued above $ 5,000. Not surprisingly traders prefer dealing with Tanzanian Customs officials and routinely divide their shipments into smaller consignments or grossly under-declare values (claiming machinery is reconditioned, for example). This arrangement differs from the franchise bidding I have described in the sense that a tax concession relies on the honesty of the auctioneer (to award to the highest bidder and to enforce the contract) while the “customs” solution relies on the honesty of the sub-contractor (for example, to collect and remit as much as possible).

22 Internal like New Zealand (Schick, 1998), or external as in the case of tax concessions.
system is corrupt, decisions to change it, and reforms implemented to counter it are all subject to the underlying corruption.

While this paper has focussed on poor tax compliance, quantifying leakages between taxpayer and tax collector constitutes only the “sexier” side of detecting corruption in Tanzania. There is a long chain between when taxes are paid and services are supplied. At each stage money changes hands, revenues in Tanzania are subject to embezzlement, in many cases through forged receipts. In a low trust world, where honesty is at a premium, the proliferation of transactions (common to bureaucracies, especially corrupt ones) is very costly to monitor.

To reassert my initial caution, I am unsure whether these results would hold in other countries and in other contexts. After all, Tanzania is an extreme case, one where social capital is low, morals have conspicuously eroded, and government has historically been viewed with much skepticism (see Hyden, 1980, for example). However, in the extreme world I have imagined, the financially constrained government disappears (and with it, the innumerable requests to donors for more training, better work conditions, and the like). Imagine a government in hiding with no historical pretense, one with police powers, but no history of revenue collection or service provision. By legally declaring the existence of a tax this government creates a monopoly right to extract wealth from the populace (here transparency and democracy probably count for something). This right can be subsequently and competitively auctioned to the highest bidder. On the expenditure side this government merely decides what to do; through competitive bidding, services like the construction of roads and the provision of education are awarded to private concessionaires. Although this type of thinking may force African countries back towards “colonialism” (since many of the most qualified sub-contractors may be foreign) I believe a smaller, decentralized, less ambitious government may be more appropriate where corruption has become entrenched. Since transaction costs change over time, many of the problems I have discussed in this paper are bound to be transitory. Franchise bidding has the advantage that it can be used experimentally, selectively, and on a temporary basis.

References


---

23 In Tanzania, a variant of this was done by the Dar es Salaam City Council who initiated fees for parking, to be collected by a private company.


Kobb, Daniel and George Mpelumbe. 1997. “Revenue Collection in five Districts in Northern Tanzania (Korogwe, Mwanga, Muheza, Lushoto, Handeni): a Summary.” Village Development Project (VDP) and Natural Resources and Buffer Zone Management Programme (NRBZ), Tanga and Lushoto Tanzania.


Chapter 3
Value Added Tax Evasion, Auditing and Transactions Matching*

Arindam Das-Gupta and Ira N. Gang

Abstract. This paper extends the standard theoretical model of tax enforcement by allowing for the cross-matching of transactions in addition to the auditing of taxpayers. For the Value Added Tax (VAT) the matching of purchase and sales invoices is an important enforcement technique. The paper examines the impact these activities on VAT enforcement and on efficiency. Transactions matching is shown to have very different compliance and efficiency effects compared to auditing: even if auditing by itself cannot induce non-zero taxpayer reports and has a low expected success rate, sufficiently intensive cross-matching can induce truthful reporting. Matching, however, may distort purchase and sales decisions. It can also cause distorted input use and output decisions even if auditing alone has no adverse effects. Conditions under which the VAT leaves input prices undistorted are studied and the content of the often-made claim, that a VAT is self-enforcing, is explored. The ability of the tax administration to enforce compliance with the VAT is shown to be sensitive to the knowledge that the tax administration has about production technology.

I. Introduction

Theoretical papers on tax evasion, beginning with Allingham and Sandmo (1972), have focused on only one aspect of the technology of tax enforcement, namely auditing. Auditing, in these papers, is an activity that leads to the discovery of the under-reporting of the tax base of the audited taxpayer by the tax authority. An important implicit assumption, which may be termed the independent audit assumption, is that auditing of one taxpayer does not systematically throw up information that can be used to detect evasion by other taxpayers. This assumption may be justified in many situations. However, in important modern contexts, such as in the enforcement of the income-tax and the value added tax (VAT), the assumption is unrealistic. The underlying transactions structure of income creation or value addition implies that auditing of one taxpayer throws up useful information on other taxpayers. This information gain arises essentially from

* The first draft of this paper was completed while Das-Gupta was a Visiting Fulbright Scholar at Rutgers University. He gratefully acknowledges support from the Fulbright Scholarship and the hospitality of Rutgers University. We also wish to thank John McLaren, Doug Marcouiller, Dilip Mookherjee, Chung Gool Moon, and Chuck Romeo for helpful comments and discussions.

1 Reviews are in Cowell (1990) and the Special Supplement to Public Finance/Finances Publiques, 1994.

2 For example, enforcement of property taxes, the retail sales tax, land taxes and import duties.
matching the receipts (for example from sales) with the expenditures (for example on purchases) of different taxpayers.

Models that neglect this interdependence run the risk of drawing misleading conclusions with regard to the effectiveness of enforcement and the efficiency or equity implications of different enforcement strategies. The major objective of this paper is to extend the standard treatment of tax enforcement through independent audits by including transactions matching, an enforcement activity which systematically throw up information of use in examining other taxpayers. An extended model is then used to examine tax enforcement effectiveness and its efficiency implications.

For concreteness we focus on the VAT, so that the information structure can be carefully specified. Furthermore, the opportunity for tax evasion under a VAT may be crucially affected by transactions matching.3

VAT evasion is also of independent interest since it has replaced other forms of sales taxes in many countries of the world in recent decades and continues to attract new converts. This popularity is partly because, in comparison with other sales taxes, a VAT is thought to have two important advantages.4 First, intermediate goods are supposed to bear no net tax. For example, under the widely used invoice method of administering the VAT, tax paid by intermediate goods producers is rebated to final goods producers against purchase invoices. As a result, marginal conditions for production efficiency are undistorted by the VAT. As is well known since the work of Diamond and Mirlees (1971), zero taxation of intermediate inputs forms part of an optimal commodity tax mix under very general conditions in a second-best world. That a consumption type VAT, as well as a retail sales tax, leads to zero taxation of intermediate goods in a second best world is an important advantage claimed for it over other forms of the sales tax. Does this advantage survive in a third best world where tax evasion is possible?5

The existence of both purchase and sales invoices for the same transaction leads to the second important advantage claimed for a VAT commonly termed "self-enforcement."6 This arises from the possibility of matching sales invoices against purchase invoices, making it difficult for intermediate goods sellers to understate sales, especially since purchasing firms have an incentive, other things equal, to declare purchase invoices to the sales tax administration (STA) and receive rebates.

Thus, a second objective of this paper is to examine the effectiveness and efficiency implications of a VAT, implemented by the widely employed invoice method, in the presence of evasion and to study the extent to which "self-enforcement" takes place. Since, of course, no tax can magically enforce itself, we must first define self-enforcement. Essentially, we identify potential self-enforcement with a situation in which the optimal voluntary declaration by taxpayers is increasing not only in the level of enforcement by the STA but also with respect to

See, for example, Sandford and Godwin 1990.
4 See, for example, Tait 1988 for a comprehensive discussion of the VAT.
5 For recent work on optimal taxation under conditions of tax evasion see, for example, Cremer and Gavhari 1994. So far as we are aware, no paper has yet examined the optimal treatment of intermediate goods in the presence of tax evasion.
6 See Sandford and Godwin, op. cit.
voluntary reports by other taxpayers. For self-enforcement, furthermore, we require that taxpayers make positive reports in equilibrium.

We construct a two industry partial equilibrium model of VAT evading firms in order to examine these questions. STA enforcement is both through auditing and cross-matching of purchase and sales invoices, the latter activity capturing the interdependence of enforcement information on different firms. Our main findings are as follows:

i. Auditing and cross-matching have very different effects on the compliance behavior of firms: Even in situations where auditing alone would lead to zero reports by firms, cross-matching can lead to zero tax evasion essentially due to the self-policing property of the VAT.

ii. Unlike auditing, cross-matching alters book-keeping and purchasing behavior of firms. This may have adverse implications for allocative efficiency in the presence of scale economies.

iii. Auditing and cross-matching have independent effects on production efficiency via input prices: Distorted input prices may obtain due to cross-matching even when auditing alone leaves input prices undistorted and vice versa.

iv. In situations where auditing alone has no effect on the output decision of firms, a problem studied by Marelli (1984), cross-matching can still affect output decisions thus leading to a violation of the product-mix efficiency conditions (i.e. Marginal Rate of Substitution = Marginal Rate of Transformation) of the economy.

The very different implications we find in comparison to a model that only allows for auditing, show that extension of the standard model to incorporate cross-matching is of importance. For the VAT, a comparison of (i) with (ii)-(iv) above suggests a possible trade-off between effective enforcement, and therefore the revenue generating ability of the VAT, and its impact on production efficiency and product-mix efficiency. Our findings also raise the possibility that the VAT is not part of an optimal commodity tax mix in a third-best world with tax evasion and costly enforcement (alternatively, if a VAT continues to be optimal, taxation of inputs must form part of a third best optimal commodity tax). This question, however, is beyond the scope of this paper.

In Section II we specify our two-industry framework in the absence of evasion. Evasion and enforcement are introduced in Section III. The specification of the information structure in Section III is intentionally detailed and based on a parable, so that the key differences between auditing and cross-matching in our model are readily apparent to the reader. In an Appendix the implications of the assumptions we make for the purchase of intermediate goods by firms and account-keeping behavior are examined. Expected utility (of profits) maximizing firms are shown to prefer splitting input purchases between all available intermediate-input sellers, leading to possible loss of scale economies. This, however, leads to a simplification of the model, which is useful for our main analysis.

The model is then specified in full and used to examine the main VAT questions, the impact of evasion and enforcement on input prices and self-enforcement. In general, the VAT is
neither self-enforcing nor leaves input prices undistorted (Section IV). Somewhat surprisingly however, it is possible for the VAT to be both self-enforcing and to leave input-prices undistorted with sufficient enforcement effort, regardless of the effectiveness of the STA in carrying out audits as measured by the (ex ante) expected success rate. Such high levels of enforcement effort may, of course, be too costly to implement. In other situations expected revenue maximizing audit and cross-matching may lead to input price distortion even if there is no difference in per unit enforcement costs across industries (Section V). Having dealt with input price distortion, in Section VI we find necessary and sufficient conditions for the VAT to be self-enforcing. The conditions are fairly undemanding, so that our paper provides theoretical support for the claim that the VAT is largely self-enforcing in the presence of cross-matching.

VAT evasion is examined in the paper under the assumption that the input-output structure of production is common knowledge. In a brief extension (Section VII), the importance of the STA's information about this structure for its ability to collect revenue from the VAT is demonstrated. Whether or not enforcement effectiveness is undermined, it is argued, depends crucially on the legal structure, specifically, if the onus of proof that an invoice is not a fake is on the taxpayer or the government.

The model in the paper is highly simplified and leaves out many features of economies and the VAT of importance in the real world. To close our discussion, limitations and extensions of our analysis are discussed (Section VIII). In particular, we argue that our simplifying assumptions with regard to market structure and also the particular procedure for cross-matching that we assume, are unlikely to qualitatively affect our conclusions as to split transactions, input distortion and self-enforcement.

II. Basic Industry Structure

The basic model we specify is very simple as the extension to cross-matching will lead to added complexity. Nevertheless, we believe that our model leaves out nothing crucial. Furthermore, even in our simple model important differences arise between the effects of different enforcement activities.

There is a final goods industry producing a homogenous good $F$ and an intermediate goods industry producing a homogenous input, $I$. Final and intermediate goods producing firms are referred to as F-firms and I-firms respectively. The intermediate goods industry is thus assumed to be perfectly competitive with many identical I-firms. The number of I-firms is determined endogenously. The final goods industry is assumed to consist of $n$ identical firms, each of whom is a monopolist in its own region, where regions are assumed to be non-overlapping and to possess identical downward sloping demand curves for the final good. The revenue of a representative F-firm is denoted $R(F)$. F-firms may buy inputs from more than one I-firm, the number of I-firms they purchase from being determined endogenously. Likewise, I-firms may sell inputs to more than one F-firm.

The cost of production for a representative I-firm is $W(I)$. $W(I)$ has a U-shaped average cost curve reaching a minimum at $I^\ast$. The cost of production for the representative F-firm consists of two parts, the cost of primary inputs, $C(F)$, and the cost of intermediate inputs, $\alpha wF$, where $w$ is the price of intermediate inputs.
where \( \alpha \), the input-output coefficient, is assumed to be constant and \( w \) is the price per unit of input purchased (and \( \alpha F \) units of input are purchased).\(^7\) The marginal primary cost function is assumed to be positive and increasing. Revenue and cost functions are assumed to be at least twice differentiable. It is assumed that no inventories are held by F-firms or I-firms: all intermediate purchases are used up in production in the same period and all final and intermediate goods produced are sold.\(^8\) Consequently, in the absence of taxation, profits are \( \pi_I(I) = wI - W(I) \) for I-firms and \( \pi_F(F) = R(F) - C(F) - \alpha wF \) for F-firms. Given competition, \( \pi_I = 0 \) and \( I = I^* \) in long-run equilibrium.\(^9\) Without loss of generality, units of input are chosen so that \( \alpha = 1 \). This is a partial equilibrium model as factor supplies underlying cost curves and the demand for final goods are taken as given.

We assume throughout the paper that all firms in a taxed industry are liable to pay sales taxes - there are no exempt firms.\(^{10}\) The ad valorem rates of tax across industries are taken to be identical and denoted by \( t \). Clearly, the average cost curves of I-firms are shifted up by a constant amount in the presence of the VAT. This implies that zero profit equilibrium will continue to occur at a per I-firm production of \( I^* \). The equilibrium price of the intermediate good under the VAT is determined by the relation \( w_*(1-t) = w \). The number of I-firms in equilibrium is denoted \( m \), where \( m \) is determined by equating demand for intermediate goods, \( nF \) with supply \( mI^* \).

Profits of a representative F-firm under the VAT are
\[
\pi(F) = R(F)(1-t) - C(F) - wF(1-t).
\]
In (1) the price of intermediate goods is reduced by the tax \( w_*, t \) rebated to producers per unit of intermediate goods purchased (against purchase invoices), to offset the tax paid by I-firms on their sales. Consequently, the net of rebate price of inputs continues to be \( w \).

III. Incorporating VAT Evasion and Enforcement Activity

Tax evasion by I-firms takes the form of under-reporting sales to F-firms. F-firms can over-report purchases or under-report sales. Penalties for evasion and the information structure are now specified. The main innovation is in the specification of the structure of cross-matching with other assumptions being similar to those in earlier work. In modeling evasion behavior, we first ensure that penalties on over-reported purchases and under-reported sales do not, by themselves, lead to differing incentives to engage in these two activities.

\(^7\)The extension of the analysis to include capital purchases requires an explicitly dynamic model as treatment of capital purchases may differ under different variants of a VAT. Either primary or intermediate inputs may, however, be interpreted as including the cost of capital services, depending on which variant of the VAT is being examined, without affecting the analysis. For a discussion of VAT variants see, for example, Due and Friedlaender (1973) and Due (1988, Chapter 16).

\(^8\)This assumption closes off one possible channel of VAT evasion discussed, for example, in Tait (1988).

\(^9\)We assume the existence of stable and unique equilibria for all sets of tax and enforcement parameters that arise in the course of this analysis without further comment.

\(^{10}\)In practice, various types of firms are exempt from tax leading to possible distortion in output and further enforcement problems.
A1. There is no penalty for not reporting purchases (firms may report purchases if they wish). F-firms will, of course, need to report purchases to the extent that they wish to claim VAT rebates.

A2. The penalty for tax evasion is levied on net underpaid taxes detected, at a constant rate $f>0$. This ensures that independently varying reported sales or purchases have no impact on penalties, provided total tax evaded is unaffected.

We now turn to the structure of accounts.

A3. Firms are required by law to issue and keep copies of sales invoices that bear the names of both the buyer and the seller. Tax returns need not, however, be supported by copies of invoices. These must be produced only if the firm is audited.\(^\text{11}\)

Thus, realistically enough, the STA will not be able to infer the identity of sellers of intermediate inputs to an F-firm or the identities of purchasers of inputs from a given I-firm in the absence of an audit.

A4. Invoices in books shown to the STA by an audited firm are (at least) equal in value to rebates claimed by the F-firm, or sales voluntarily declared by the I-firm, so that accounts and voluntary reports are consistent.

This assumption, which simplifies the analysis, may not be entirely innocuous.\(^\text{12}\) Finally, we make an assumption that permits us to focus on the case of I-firms and F-firms that deal with each other at arms length. In comparison with the other assumptions made here, the study of situations where this assumption does not hold may be an important task for the future.

A5. Firms deal at arms length. In particular, there is no collusion between I-firms and F-firms to conceal transactions from tax authorities.

We must now describe the information structure of the sales tax administration. The formal statement of our assumptions is followed by a description of the organizational parable we have in mind in this specification.

A6. Prices of the I-good and f-good as well as the input-output coefficient are common knowledge.

In a brief extension, we show that the STA's knowledge about the input-output coefficient is crucial to its ability to enforce VAT compliance. Knowledge of the input-output coefficient implies that the STA can establish in the appropriate court of law that no more than $\alpha F$ units of the intermediate input are required to manufacture $F$ units of output. An example is the case of a grocery store that merely acts as a regional outlet for various consumer goods. An example of a case where the STA's knowledge will be limited is a tailoring establishment: different tailors will:

\(^{11}\)See, for example Tait 1988, Chapters 13 and 14. South Korea appears to be the only country that required invoices to be sent to the tax office.

\(^{12}\)We discuss this in the concluding section.
use differing amounts of material to make similar suits for identical customers. A second example is where two different processes are in parallel use to make a product - the inefficient process not having been completely phased out.

Next, consider tax audits by the STA.

A7. Firms to be audited are all selected prior to the cross-matching process: no resources are available for a second round of audits after cross-matching. The possibility of additional audits may be important in practice. The assumption is innocuous here, given identical firms.

A8. The probability of audit for I-firms and F-firms are $q$ and $p$ respectively, $0 \leq p, q \leq 1$.

Audit probabilities are taken to be equal to the fraction of firms audited. The number of firms audited in an industry is a policy variable.

A9. With probability $e$, $0 \leq e \leq 1$, all sales come to light in an audit, while with probability $1-e$, no unreported sales are revealed by the audit. $e$ is technologically given.

This all or nothing assumption is standard in the literature as there is little to be gained from allowing for partial discovery of evasion in an audit. It will be obvious later that STA technological ability differing across industries, or differing per unit enforcement costs, would have made our task easier by enhancing the possibility of input price distortion with VAT evasion. The implicit assumption that detection probabilities are independent across firms made here is the key assumption driving purchase splitting results in the paper. This suggests that the consequent loss of economies of scale, if present, is an important component of the cost of tax enforcement in practice.

To ensure that some evasion takes place we place an upper bound on $e$, similar to the standard condition imposed on the audit probability in the literature:

A10. $e(1+f) < 1$.

Next, turn to the cross-matching of transactions. Invoices brought to light in audits form the information base for cross-matching. Over-reporting of purchases under a VAT will never be strictly optimal for F-firms if the input-output coefficient is common knowledge. Fake invoices are considered in the extension in Section VII. South Korea is, to date, the only country that has attempted to implement a mechanism to match all purchase and sales invoices that its administration was aware of (Tait 1988). Consequently, we allow for partial cross-matching of a subset of transactions on which the STA has information.

A11. A fraction $s$ ($S$) of I-firm invoices known to the STA is matched with F-firm (I-firm) reports.

The information potentially available to the STA for cross-matching under our assumptions is summarized in Table 3-1. The key difference between audited and unaudited firms arises from the information with the STA about their transactions. For unaudited forms, the STA knows only
total purchases (sales) reported but not the identity of the firms from which purchases were made (to which sales were made).

To be concrete, imagine the following procedure for matching. After auditing, invoices are sent to the STA's matching division. The matching division must sort invoices received from each I-firm according to the F-firm named in the invoice. Before these invoices are received, the F-firm's tax file contains: (a) for unaudited F-firms, the tax return; (b) for audited F-firms, the return, a record of additional sales detected on audit, and invoices of purchases from different I-firms that are revealed voluntarily or discovered by the STA. The information for cross-matching of I-firm returns is similar.

Table 3-1: Information Potentially Available to the STA for Cross-Matching

<table>
<thead>
<tr>
<th>INFORMATION ON F-firms</th>
<th>INFORMATION ON I-firms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>F-firm audited</strong></td>
<td><strong>F-firm not</strong></td>
</tr>
<tr>
<td><strong>unsuccessfully</strong></td>
<td><strong>audited</strong></td>
</tr>
<tr>
<td><strong>I-firm audited</strong></td>
<td><strong>F-firm not</strong></td>
</tr>
<tr>
<td><strong>successfully:</strong></td>
<td><strong>audited</strong></td>
</tr>
<tr>
<td>Purchases from the I-</td>
<td>Total purchases</td>
</tr>
<tr>
<td>firm reported by F-firm for matching against</td>
<td>imputed from sales reported by F-firm for matching against total reported or discovered sales by audited I-firms to the F-firm.</td>
</tr>
<tr>
<td>sales to the F-firm by I-</td>
<td></td>
</tr>
<tr>
<td>firm.</td>
<td><strong>F-firm audited</strong></td>
</tr>
<tr>
<td><strong>F-firm not</strong></td>
<td><strong>audited</strong></td>
</tr>
<tr>
<td><strong>unsuccessfully</strong></td>
<td><strong>missed</strong></td>
</tr>
<tr>
<td><strong>I-firm audited</strong></td>
<td><strong>F-firm not</strong></td>
</tr>
<tr>
<td><strong>unsuccessfully:</strong></td>
<td><strong>audited</strong></td>
</tr>
<tr>
<td>Purchases from the I-</td>
<td>Sales to F-firm reported by I-firm for matching against purchases from I-firm by F-firm.</td>
</tr>
<tr>
<td>firm reported by F-firm for matching against</td>
<td></td>
</tr>
<tr>
<td>sales to the F-firm by</td>
<td></td>
</tr>
<tr>
<td>I-firm.</td>
<td></td>
</tr>
</tbody>
</table>

This structure renders it impossible for the STA to limit in advance the sorting of invoices according to firms named in invoices: I-firm (F-firm) invoice records must be sorted even for F-firms (I-firms) that have been successfully audited. However, the STA can choose to sort records of only a subset of audited firms. Furthermore, after sorting is complete, the matching wing can select a subset of assessment files to actually carry out matching or tallying of sales and purchase totals. So assume that:
A12. Invoices of a randomly chosen subset of audited firms are sorted and all sorted records are matched.

There are thus four enforcement activities: matching of invoices for firms in either industry and auditing of firms in either industry. This structure essentially captures the informational advantage of a VAT administration over other sales tax administrations in detecting sales by intermediate goods industries. Information on tax evasion by final goods firms may, however, be common to a larger class of taxes (such as the multi-point cascade sales tax).

The implications of these assumptions for purchase, sales and book-keeping behavior of expected utility of profit maximizing, tax evading firms in the presence of a VAT administered by the invoice method are examined in the Appendix. It is shown that F-firms making intermediate purchases from as many I-firms as possible and recording equal fractions of purchases from each firm in account books weakly (strictly) dominates any other purchase and book-keeping pattern for risk neutral (risk averse) firms. Similarly, I firms prefer to record equal fractions of sales to different F-firms in their books. Consequently, we assume such purchasing and book-keeping behavior by firms.

IV. Evasion Decisions by Firms and Input Price Distortion

Our model is related to existing models of sales tax evasion by monopoly firms (Marelli, 1982) and competitive firms (Virmani, 1989). Two new factors come into play in comparison with earlier work. First, it is possible to use cross-matching of invoices to detect evasion by firms who make it through the audit round unscathed. Second, F-firms have an incentive to over-report purchases (that is, report fake purchases) in addition to under-reporting sales. Over-reporting is ruled out by the assumption that the input-output coefficient is common knowledge. The consequence of relaxing this assumption will be explored later. To simplify the analysis of evasion decisions, we follow Virmani (1989) and assume that firms are risk neutral.

Given equal recording of sales to F-firms, the expected profit of a representative I-firm is given by

$$\pi_i(I, \Phi) = w, I - W(I) - t_w, I[\Phi + \lambda(1 + f)],$$

where,

$$\lambda = qe(1 - \Phi) + q(1 - e)pS[\Theta + e(1 - \Theta) - \Phi] + (1 - q) \sum_{j=1}^{S_p} b(Sn, e, j)[\Theta S + \left(1 - \Theta\right)j - \Phi],$$

for $\Phi < \min[\Theta + e(1 - \Theta), S_p]$,

is the expected detection by the STA of under-reported sales. In (3), $J_f(\cdot)$ is the minimum number of F-firms that need to be audited to establish evasion by unaudited I-firms through cross-matching\(^{13}\) and $S$ is the fraction of audited F-firms selected for cross-matching. The first term in (3) is the expected detection of under-reporting through cross-matching. The second term gives the expected detection of under-reporting through cross-matching if the I-firm is unsuccessfully audited given that F-firms report a fraction $\Theta$ of their sales. The term is zero if $\Phi \geq \Theta + e(1 - \Theta)$. The last term is the expected detection of under-reporting through cross-matching for unaudited

\(^{13}\) This is discussed further in the Appendix.
I-firms, given that $\Phi < S_p$. $b(s_{\Phi}e, j)$ is the probability that sales to $j$ F-firms will be detected upon cross-matching given that a fraction $p$ of the $n$ F-firms are audited from which a further fraction, $S$, are selected for cross-matching. $b(.)$, it is argued in the appendix, is a binomial distribution where $p e$ is the probability of a successful audit.

Similarly, for F-firms, expected profits are given by
\[
\pi(F, \Theta) = R(F) - C(F) - w_v F - [\Theta + \Omega(1 + f)][R(F) - w_v F],
\] where
\[
\Omega = pe(1 - \Theta) + p(1 - e)qs[\Phi + (1 - \Phi)e - \Theta]
+ (1 - p) \sum_{j(m, \Phi, \Theta)} b(s_{\Phi}q, e, j)[\Phi sq + \frac{(1 - \Phi)j}{m}] \theta \text{ for } \Theta < \min[\Phi + e(1 - \Phi), sq].
\] (5)

The interpretation of the three terms in (5) is similar to the case of I-firms.

From (2) and (3) the fraction of sales that will be reported by an I-firm is independent of the level of output. Consequently, given any report, the average cost curve will shift up vertically by the amount of the effective tax (i.e., the expected tax cum penalty) in comparison with the no tax cost curve. Thus, the long run output of an I-firm will still be $I^*$. The number of I-firms will continue to be determined by the market equilibrium condition
\[
n_F = mI^*
\] (6)

Equations (2) to (6) bring together the equations of the model being analyzed in this section. From (3) and (5) it can be seen that the fraction of sales reported by I-firms and F-firms are mutually interdependent. Consequently, equilibrium requires, in addition to mutually consistent profit maximizing output decisions, equilibrium expectations concerning the reported sales fractions, $\Phi$ and $\Theta$. This can be found by solving "reaction functions" of representative I- and F-firms for $\Phi$ and $\Theta$. Equilibria are further examined in Section VII below.

We now define self-enforcement. The VAT will be said be potentially self-enforcing if the optimal report of F-firms is increasing in the optimal report of I-firms and vice-versa for a given set of enforcement parameters. If, in addition, firms report positive sales in equilibrium, the VAT will be said to be self-enforcing. The latter requirement ensures that reports by firms do, in fact, influence reports by other firms rather than merely having the potential to do so. We now turn to an examination of conditions under which input prices remain undistorted. It is clear that for input prices not to be distorted by the VAT, I-goods must bear no net tax and receive no net subsidy. Our concern here is, therefore, in what enforcement regimes this situation obtains. We argue, first, that the model has the following, rather surprising, property.

**Proposition 1:** For any value of the efficiency parameter $e$, $0 < e < 1$, the STA can ensure that firms report their sales truthfully with sufficiently intensive cross-matching and auditing, due to the self-enforcement property of the VAT.

Note that, given the assumption $e(1 + f) < 1$, both I-firms and F-firms will make zero reports if there is no cross-matching since, in this case, expected profits for both types of firms are

---

This differs from the result of Virmani 1989 who concludes that tax evading competitive firms will produce below the minimum efficient scale. This is because he assumes that concealment of sales is costly with costs increasing with the proportion of sales the firm attempts to conceal from tax auditors.
decreasing in the fraction of sales reported regardless of audit rates. Thus positive voluntary reports by firms, if they obtain, must be due entirely to the additional effect of cross-matching.

To prove the proposition, suppose, initially, that cross-matching of unaudited firms is absent, so that the third terms in (3) and (5) drop out. Differentiate (2) with respect to \( \Phi \) using (3). The derivative is \((1+f)[q(e + q(1-e)pS)]\)-1. Clearly, for sufficiently large \( q, p \) and \( S \) this will be positive. In such a case expected profit maximization will require that \( I \)-firms make the report \( \Theta=\Phi+e(1-\Phi) \) so that no evasion can be detected through matching. They will not report any higher amount since auditing alone is unable to deter under-reporting. Similarly, with sufficiently intensive auditing and cross-matching \( F \)-firms will report \( \Theta=\Phi+e(1-\Phi) \) from (4) and (5). But, these reaction functions intersect only at \( \Theta=\Phi=1 \) which, therefore, must hold in equilibrium! Clearly, the argument will go through even if we allow, in addition, cross-matching of unaudited firms. This completes the argument. The precise condition for this full compliance equilibrium is, of course,

\[
\min[qe + q(1-e)pS, pe + p(1-e)qs](1 + f) \geq 1. \tag{7}
\]

The rather unexpected interpretation of this finding is that large-scale cross-matching, possibly with the aid of high speed computers, can compensate to a large extent for lack of ability to detect evasion in traditional audits. Of course, large scale matching may be too costly to implement, though the cost should be greatly reduced by the use of high-speed computers.

Now consider the opposite case where auditing and matching rates are low enough that both types of firms make zero reports. From (2)-(5) a sufficient condition for zero reports to be optimal for both types of firms is

\[
\max[qe + q(1-e)Sp + (1-q), pe + p(1-e)sq + (1-p)](1 + f) < 1. \tag{8}
\]

This follows since zero is the optimal response of each type of firm, \textit{given} a zero report by the other type if the condition holds. With zero reports by both \( I \)-firms and \( F \)-firms, (2) and (4) reduce to

\[
\pi(I, 0) = w_v I - W(I) - t_w e[q + (1 - qe)Sp](1 + f), \tag{9}
\]

and

\[
\pi(F, 0) = R(F) - C(F) - w_v F - te[p + (1 - ep)sq](1 + f)[R(F) - w_v F]. \tag{10}
\]

In long run equilibrium, the price if \( I \)-goods can be determined from (9) to be

\[
w_v = \frac{w}{1 - te(1 + f)[q + (1 - eq)Sp]}. \tag{11}
\]

If \( S=s=1 \) or \( q=p \) and \( S=s \), the denominator of (11) is identical to the net of rebate cost of inputs in (10) implying that input costs continue to be undistorted in the presence of evasion when enforcement effort is sufficiently weak (that is, (8) holds) and there is complete cross-matching.

A third case of undistorted input prices results from equal reported sales fractions for \( I \)-firms and \( F \)-firms. This can be found by inspection of (2)-(3) and (4)-(5) to be where \( S=s, p=q \) and \( n=m \) in equilibrium. We have thus shown that
Proposition 2: Sufficient conditions for a VAT to leave input prices undistorted are that

[i] \((11)\) holds; or

[ii] either \(S=s=1\) or \(q=p\), \(S=s\) and, furthermore, \((8)\) holds; or

[iii] \(S=s\), \(p=q\) and \(n=m\).

Self-enforcement also occurs in case [i].

Intuitively, the proposition identifies different cases in which the effective tax rate on I-firms and the per unit effective rebate on intermediate goods purchases by F-firms are equal even in the presence of evasion. The very special nature of these conditions shows that undistorted input prices can be expected to be the exception rather than the rule.

V. Revenue Maximizing Allocation of Audit Resources and Input Price Distortion

Our concern in this section is to demonstrate that STA revenue maximization can require that audit probabilities and matching rates differ across industries even if there is no difference in the cost of auditing or cross-matching across industries. As discussed above, this will strongly imply a possible conflict between the revenue goal and input-price neutrality. Furthermore, the analysis enables us to shed some light on factors that should influence the design of audit and matching strategy. To carry out the analysis, we assume that the STA is able to commit to an audit-cum-matching policy. A principal-agent analysis is then appropriate. We restrict attention to the case where audit costs are high enough so that \((8)\) holds and firms find it optimal to make zero reports. Since a non-zero level of enforcement requires some auditing our examination can be restricted to checking if, at interior optima, \(S=s\) and \(p=q\) hold or if, at the optimum, \(S=s=1\).

The equation for government revenue, \(G\), is given by

\[
G = n[R(F) - w_v,F]t_v + nw_v,Ft_v - (np + mq)g_a - (Sp + sq)nmg_M, \tag{12}
\]

where \(t_v\) is the effective tax rate on I-firms equal to \(te(I+f)[q+Sp(1-ep)]\) (so that \(w_v = w/(1-t_v)\); \(t_v\) is the effective tax rate on F-firms equal to \(te(1+f)[p+sq(1-ep)]\); \(q=A/m\) and \(g_a\) and \(g_M\) are respectively the cost per audit and per matched transaction assumed to be identical across industries and constant per audit. The equations determining \(m\) and \(F\) are \((6)\) and the first order condition derived from \((10)\), respectively.

The first order necessary conditions for revenue maximization are

\[
\begin{align*}
(1 - qe)\xi_1 - nmg_M &= 0 \\
(1 - pe)\xi_2 - nmg_M &= 0 \\
\left(\frac{s}{q} \frac{\partial G}{\partial S} \right) + (1 - Spe)\xi_1 - mg_a &= 0 \\
\left(\frac{s}{p} \frac{\partial G}{\partial S} \right) + (1 - sqe)\xi_2 - ng_a &= 0
\end{align*}
\]

The equation for government revenue, \(G\), is given by

\[
G = n[R(F) - w_v,F]t_v + nw_v,Ft_v - (np + mq)g_a - (Sp + sq)nmg_M, \tag{12}
\]
where,
\[
\xi_1 = te(1 + f)w_nF \left[ \frac{(1-t_e)G_F}{D} \right],
\]
\[
\xi_2 = te(1 + f)\left[ n(R(F) + w,v) + \frac{(R'(F) - w,v)G_F}{D} \right],
\]
\[
D = (1 - t_e)R^* - C^* + te(1 + f)qF \left[ (1 - Spe)^{1-t_e} + (R'(F) - w,v)s(1-ep) \right] < 0.
\]

(14)

In (14), \( G_F \) is the partial derivative of \( G \) with respect to \( F \). The negative sign of \( D \) follows if the equilibrium in the 2 market model is stable. From (13), \( p=q \) and \( S=s \) will be a solution if and only if \( n=m \) and \( \xi_1=\xi_2 \). A necessary condition for the solution \( S=s=1 \) can be found to be \( g_A/g_m > (nm)^{0.5} \). Thus audit and matching rates depend on there being an equal number of firms in the two industries besides appropriate cost and demand conditions. They will not, in general, leave input prices undistorted.

VI. Self-Enforcement

One case in which the VAT is self-enforcing, where (7) holds, has been considered in section IV. Likewise, one case where the VAT is not self-enforcing has been found to be where (8) holds. Are more general results available? In fact a condition related to (8) characterizes cases of equilibrium self-enforcement under the VAT. The condition merely ensures that all-firms in the economy make non-zero reports, a requirement that appears to be relatively undemanding. The analysis is, however, complicated by the fact that optimal reported output fractions will take on only a discrete set of values corresponding to the "kink points" of (3) and (5). To simplify the analysis assume that the number of I-firms in equilibrium and the number of F-firms are large enough so that a differentiable approximation to the binomial distribution \( b(.) \), to be denoted \( \beta(.) \), can be employed. In this case, (3) and (5) are replaced by:

\[
\lambda = qe(1 - \Phi) + q(1 - e)pS[\Theta + e(1 - \Theta) - \Phi]
+ (1 - q) \int_{J(n,\Theta,\Phi)} b(Snp,e,j)[\Theta Sp + (1 - \Theta)j - \Phi] dj,
\]

for \( \Phi < \min[\Theta + e(1 - \Theta), S_p] \) \hspace{1cm} (15)

and

\[
\Omega = pe(1 - \Theta) + p(1 - e)qs[\Phi + (1 - \Phi)e(1 - \Theta)]
+ (1 - p) \int_{J(m,\Theta,\Phi)} b(sA,e,j)[\Theta sq + (1 - \Phi)j - \Theta] dj\]

for \( \Theta < \min[\Phi + e(1 - \Phi), S_q] \). \hspace{1cm} (16)

In (15) and (16), \( J_I \) and \( J \) are no longer integers but are given by \( J_I = n(\Phi - \Theta S_p)/(1 - \Theta) \) and \( J = m(\Theta - \Theta S_q)/(1 - \Phi) \). With this simplification we can state:

**Proposition 3:** Suppose that the expected profits of I-firms are given by (2) and (15) while the expected profits of F-firms are given by (4) and (16). Then the VAT is self-enforcing if and only if

\[
\min \left[ qe + q(1-e)S_p + (1-q), pe + p(1-e)S_q + (1-p)(1+f) \right] > 1.
\]

(17)
The condition in the proposition merely requires non-zero sales reports by at least one kind of firm. In other words, the VAT with cross-matching, is self-enforcing if at least one type of firm independently decides to make a positive sales report, no matter how low this is. The interdependence of firms induced by cross-matching then takes over. Thus self-enforcement due to cross-matching, the model predicts, should be the exception rather than the rule.

To prove the proposition we first show that (17) is necessary and sufficient for non-zero reports by both kinds of firms. We next argue that reports of each type of firm are increasing in the reports of the other, whenever both have interior solutions. Finally, we describe the set of interior equilibria. We provide the argument underlying Proposition 3 for F-firms, noting that an analogous argument holds for I-firms. Differentiating (4) with respect to \( \theta \) using (16) and the expression for \( J(.) \) yields the expression in (18):

\[
-1 + (1 + f)[pe + p(1 - e)qs + (1 - p) \int_{\theta(0, \theta)}^{\theta} \beta(s, e, j) dj].
\]

It is easily seen from (17) that a necessary and sufficient condition for a zero report to be optimal for F-firms is that 

\[
(1 + f)(pe + p(1 - e)qs + (1 - p)) \leq 1.
\]

An analogous condition holds for I-firms. Consequently, (17) is necessary and sufficient for both I- and F-firms to make positive reports. Furthermore, setting the expression in (18) to zero, we see that this implies \( J(.) \) being constant regardless of \( \Phi \). From the expression for \( J(.) \) above, this implies that \( \Phi \) is increasing with \( \Phi \). Consequently, the second leg of the proof is complete if self-enforcement holds in the case of there being no interior solution to (18). With no interior solution, we will have \( \phi \geq sq \), at which values the cross-matching of unaudited F-firms cannot detect additional evasion. Second, from (18) it is also the case that 

\[
(1 + f)(pe + p(1 - e)qs) \geq 1.
\]

In this case, as seen in Section IV above, it is optimal for the F-firm to set \( \phi = e(1 - \Phi) \). Once more, \( \Phi \) is increasing in \( \Phi \) and \( \Phi \) will be positive under (18). This completes the argument.

The nature of self-enforcement equilibria, at the intersection of reported sales reaction functions of I- and F-firms is shown for the case where 

\[
(1 + f)\max[pe + p(1 - e)qs, qe + q(1 - e)pS] < 1
\]

in Figure 1. The first condition ensures that, at the optimum, \( \phi \leq sq \) \( \phi \leq Sp \). Two possible cases arise for each type of firm, corresponding to \( sq \leq e \) and \( sq > e \) for F-firms and \( Sp \leq e \) and \( Sp > e \) for I-firms. From (18), furthermore, reaction functions of F-firms are given by \( J = \text{constant} \). An analogous reaction function can be found for I-firms. Consequently, the reaction functions are straight lines. The discontinuity occurs at the point where cross-matching of audited firms starts to have an effect or when \( sq = e(1 - \Phi) \). Self-enforcement is illustrated by the fact that the reaction functions are positively sloped so that higher reports by one kind of firm induce higher reports by the other type of firm. The other equilibrium with self-enforcement, analyzed in Section V, is the point marked e in the diagram.
VII. Extension to Imperfectly Known Technology

Our final demonstration is to show that effective cross-matching is important for the VAT. Allowing for fake invoices, which vitiates the effectiveness of cross-matching, may seriously affect VAT revenue performance. Specifically, we show that it is possible for F-firms to reduce their tax liability to zero (or less) by optimally over-reporting purchases if we relax the assumption that the STA knows the true input-output coefficient and, furthermore, if the onus of proof that an invoice is fake is on the STA. The question of onus of proof arises when an invoice that cannot be matched against any duplicate invoice from the other party named in the invoice. If the invoice itself is prima facie evidence of payment, then, in the absence of other evidence, the STA will have to bear the revenue loss. If, however, it is the taxpayer who must provide additional evidence of genuineness when a matching document cannot be found, then revenue loss from fake receipts will be curtailed by matching. Thus, under the assumptions above, when a fake invoice naming a particular I-firm is submitted by an F-firm, the STA may be unable to determine whether the F-firm is making a fake rebate claim or the I-firm is suppressing its sales unless, of course one of the firms has been successfully audited.

Assume that the STA knows only that the true input-output coefficient is less than $\alpha^*, \alpha^* > 1$. Furthermore, let the ratio of purchase invoices submitted to the STA to true invoices be $\mu > 1$.

---

15 Inflated claims of VAT refunds are a serious problem in practice. Tait (1988) points out that 44 percent of all VAT fraud in the Netherlands had to do with inflated refund claims. He points out that "businesses have been established solely to invent and print false invoices for sale to those wishing to defraud the revenue" (p 307). Furthermore, if capital purchases are allowed for and the VAT component of the cost capital goods qualifies for rebate, the problem can be much more serious.

16 The device of naming a fictitious firm is precluded in most real world VATs as rebates are limited to invoices from intermediate goods suppliers registered with the STA. See Tait (1988) for a discussion.
1. Since the total number of purchase invoices exceeds the actual number of invoices that can be verified through cross-matching, cross-matching will play no role in detecting over-invoicing by F-firms. Consequently, the expected payments to the STA by the F-firm will be given by

\[ t{\Theta R(F) - \mu_w,F + pe[R(F)(1 - \Theta) - w,F(F(1 - \mu))(1 + f)].} \] (19)

Since expected taxes can be seen to be increasing in \( \Theta \) (given \( e(1+f)<1 \), it will be optimal to set the reported sales fraction to \( \mu^* \). Making this substitution shows that expected taxes will be negative if \( R(F)/F < \alpha^*w \). That is, the STA has to be unable to disprove that the cost per unit of intermediate inputs exceeds the price of the final goods. In this event, F-firms will over-report sales and pay less per unit of output than the rebate they receive. Clearly, while this may be possible for some firms in some periods, it cannot be considered an endemic shortcoming of the VAT. More generally, of course, greater evasion is facilitated if firms can inflate refund claims since not only can sales be under-reported but net-of-rebate tax paid per unit of output will be less than in the case of known technology. Consequently, the case of known technology that we studied above, under which cross-matching works well, is the case in which VAT enforcement has its greatest relative effectiveness.

**VIII. Conclusions and Limitations**

We have found that cross-matching can distort purchase and output decisions. The industry structure of the model used in the analysis lacks generality. So do some elements of the cross-matching procedure the STA is assumed to adopt. To what extent are our results likely to withstand generalization or modification? We believe generalizations in several directions can be accommodated.

Consider, first, the industry structure. The result on purchase splitting in the Appendix is independent of industry structure and so will be unaffected if this is altered. The analysis of input price distortion depends essentially on their being a net tax or subsidy on intermediate industries. Once again, industry structure has only a tangential role. The analysis could equally well have been for a competitive or monopolistically competitive F-industry at the cost of having to take account of long-run zero profit conditions. An interesting possibility thrown up by the analysis in the paper is the incentive firms have to vertically integrate if transactions matching leads to loss of economies of scale.\(^{17}\) An exploration of this issue must await further work.

Regarding the information structure for cross-matching, two assumptions need to be questioned. The first is assumption A4 that voluntary reports are reflected in the books maintained by firms. For I-firms, maintenance of accounts may not, in practice, be required in all countries. On the other hand, some countries impose a penalty on firms that do not maintain proper accounts. For F-firms, clearly, rebates claimed will be disallowed if not supported by invoices. The ability of the STA to use cross-matching to ensure compliance by I-firms is reduced if accounts do not support voluntary declarations. In such a situation self-enforcement will be adversely affected. Self-enforcement and also the ability of the STA to effectively use cross-matching will also be effected adversely with multiple, substitutable, intermediate inputs.

\(^{17}\) We are indebted to Dilip Mookherjee for this point.
A second problem may be the sequence of actions by the STA in auditing and cross-matching that form part of our parable on cross-matching procedure. Other organizational procedures can lead to differences in the particular form of the equations of the model. Nevertheless, two insights from our analysis appear to have some claim to generality. The first is that more purchase or sales transactions makes matching difficult so that the incentive to split purchases that we found is not merely the result of our matching procedure. This is borne out, for example, by the analysis in Das-Gupta (1994) who shows that compliance is adversely affected in the presence of transactions splitting. The second insight is the interdependence of voluntary reports by firms that is likely to persist provided there is some cross-matching. It is, however, worth examining alternative organizational arrangements explicitly in order to identify least cost organizational arrangements and distortionary consequences.

Despite the special structure of the model, we have been compelled to consider aspects of tax administration not usually dealt with in theoretical models of tax evasion. For example, the demonstration in Section V involves studying the allocation of STA resources across enforcement activities, a dimension of enforcement strategy that has not received attention in the literature. Other aspects of administration are still treated as "black boxes". These aspects relate, firstly, to the actual methods of detection of tax evasion on audit and secondly, to the STA's system for identifying taxpayers. Since the extensions we did incorporate have enabled a richer analysis of administrative activity, for example the problem of allocation of STA resources, further extensions should be worthwhile. Other aspects of tax administration include, for example, collection lags in an inflationary environment, policy for registration of firms and legal requirements or associated non-compliance penalties for maintenance of books of account.

The modeling of the VAT in this paper is also simplified and leaves out many features of a real world VAT. This is particularly true in the context of open economies and when considering investment decisions. In consequence, while the paper does call into question oft made claims about the efficiency of a VAT, it is still possible that the VAT remains superior to other types of sales taxes. Clearly, much additional work needs to be done in analyzing the VAT and sales taxes in general. A re-examination of the optimal commodity tax question, to be of some utility, must await a clarification of the major unresolved issues.

\[\text{Bagchi, Bird and Das-Gupta (1995) and Slemrod and Yitzhaki (1996) discuss tax administration issues.}\]
References


Appendix: Consequences of Cross-Matching
For Account-Keeping, Purchase and Sales Transactions

Here three questions are answered. What fraction of purchases from an I-firm will the F-firm record in the accounts it shows to the STA, given the fraction of total sales it voluntarily declares? How much will an F-firm purchase from each I-firm given the number of I-firms it deals with and its total planned sales? How many I-firms will the F-firm make purchases from? For each of the three decisions, risk-neutral firms will act so as to minimize the expected detection of unreported sales.

Consider a representative F-firm. Since the input-output coefficient is normalized to unity, the quantity of input purchased is equal to the quantity of output sold. For this section total purchases are also normalized to one unit. The fraction of actual sales reported to the STA by an F-firm is denoted by $\theta$. Since the input-output coefficient is known, the STA can infer that at least $\theta$ units have been purchased by the F-firm.

First examine the purchase by the F-firm from each I-firm and the proportion it records in its books of account taking as given the total number of I-firms, $z$, from which the F-firm makes purchases. The intuition here is strong. Given the fraction of sales reported by the F-firm, suppose that $k$ of the I-firms from whom purchases are made are audited. Since any combination of $k$ out of the $z$ I-firms can be picked with equal probability under random auditing, roughly equal quantities should be purchased from each firm and a constant fraction, $\theta$, of these purchases should be recorded in the books shown to the STA if the F-firm is audited. With purchases of unequal size, the number of I-firms that have to be audited and matched to detect under-reporting could fall below that with equal sized purchases making the detection of under-reporting more likely. Of course, the number of I-firms that have to be audited and matched to detect evasion by the F-firm will depend on the fraction of sales reported by I-firms (denoted $\phi$). We state the result as a Lemma.

**Lemma:** If a risk-neutral F-firm under-reports its sales to the STA, then for any set of values of (a) the total sales of an F-firm, (b) the proportion of sales it reports to the STA and (c) the number of I-firms it makes purchases from, no other pattern of purchases and bookkeeping rules leads to lower expected detection of under-reporting than equal purchases from all I-firms and equal amounts of each purchase recorded in the F-firm's books of account.

**Proof of the Lemma:** The result depends on showing that equal purchase and accounted fractions belongs to the set of undominated purchasing and accounting strategies. Note that the average purchase by an F-firm from an I-firm is $1/z$. Consider, first, the case where the F-firm has not been audited. For such firms, books of accounts remain unexamined and are irrelevant. Given random auditing, if $k$ I-firms ($k > z\theta$) from which purchases have been made are audited and selected for matching, then the expected fraction of total purchases by the F-firm that will be identified by the STA is $k\theta/z$ regardless of the exact pattern of purchases from different I-firms. Again, since the probability of unreported sales by an I-firm being detected is a constant fraction $e$, (and thus independent of the probability that under-reported sales by any other I-firm are detected), the expected value of additional sales to the F-firm discovered from these $k$ I-firms is $ke(1-\phi)/z$ regardless of its exact pattern of purchases. Adding the two parts gives the total.
expected discovery of unreported purchases when \( k \) I-firms are audited and matched as \( k[\Phi + e(1-\Phi)]/z - \Theta \) independent of the pattern of purchases.

Now let \( J \) be the minimum number of I-firms that have to be audited and matched for under-reporting by the F-firm to be detected, given that equal amounts are purchased from all I-firms. Clearly, \( J \) is the smallest integer such that \( J \geq \Theta z \). If \( J-1 \) I-firms are audited and matched, then no under-reporting will be detected under equal purchases. However, with certain other purchase patterns, \( J-1 \) transactions may be sufficient to detect under-reporting - such patterns will therefore be sub-optimal. Equally clearly, no unequal purchase pattern can give rise to a situation where more than \( J \) I-firms are required to detect evasion. Hence the Lemma is trivially true for risk neutral firms. Now examine the case of unsuccessfully audited F-firms. Note that exactly a fraction \( \Theta \) of total purchases is recorded in the books of account of the F-firm (the firm will pay additional taxes if more is recorded and less is ruled out by assumption). The expected discovery from each I-firm is now \( [\Phi + e(1-\Phi) - \Theta_j]x_j/z \), where \( x_j/z \) is the actual amount purchased from the \( j \)th I-firm and \( \Theta_j, z+1\Theta_jx_1+..+z\Theta_jx_z=\Theta_j \) is the amount of this purchase recorded in the F-firm's books of account. Once again, the expected discovery if \( k \) I-firms are selected is independent of the exact pattern of purchases being equal to \( k[\Phi + e(1-\Phi) - \Theta]z \), so that equal purchases and accounts will do as well as any other pattern.

**Remark:** Bounds on optimal purchase patterns can be obtained as follows. Denote the maximum quantity purchased from any one I-firm by \( \tau \) and the corresponding minimum purchase quantity by \( \tau_m \). Then it must be true that \( \tau + [(J-2)(1-\tau))/(z-1)] \leq \Theta \leq J/z \). This follows from the restriction to \( J/z \) on the combined purchases from \( J-1 \) I-firms and the fact that \( \tau \) will be at its maximum if all other transactions are as small as possible since any combination of them can be selected for audit. Since there are \( z-1 \) other I-firms that have made sales to the F-firm all other purchases must be equal to \( (1-\tau)/(z-1) \). The left hand inequality states that the total detection given that \( J-1 \) I-firms are audited and matched must not exceed \( \Theta \). Rearrangement gives \( \tau \leq [(2z-J)/z(z+1-J)] \) and \( \tau_m \geq (z-J)/(z+1-J) \). This upper bound on \( \tau \) is increasing with \( J \) and does not impose any restriction on the size of \( \tau \) if \( J=z \) but yields an interior bound if \( J<z \). That is if all \( z \) I-firms have to be audited to detect evasion, \( \tau \) can be any number less than \( \Theta \). At the other extreme, if \( J=2 \), \( z \) must be at least 2 for detection to be possible with fewer transactions) then \( \tau = 2/z \).

Other patterns of purchases and accounts may do just as well as equal purchases, provided amounts are not too different, but cannot do better. For risk averse F-firms, unequal purchases lead to increased risk\(^9\) without lowering expected detection of under-reporting. Consequently, risk averse firms will strictly prefer equal purchases and recorded amounts. With regard to sales by I-firms, a parallel argument shows that sales to F-firms of equal size and a constant fraction of each sale recorded in the I-firm's books dominates other patterns of sales with cross-matching. Consequently, without any real loss, we assume from here on that all purchases by F-firms are of equal size and that equal amounts of each purchase (sale) are recorded in books of account by F-firms (I-firms). This assumption - and the arguments justifying it - is our answer to the first two questions posed at the beginning of the Appendix.

---

\(^9\) Increasing risk in the sense of a mean-preserving spread.
Now turn to the third question, the number of I-firms from which an F-firm makes purchases, given equal sized purchases. The answer is not obvious \textit{a priori}, since there is a trade-off between a lower probability of discovery and a higher fraction discovered from each I-firm if purchases are made from fewer I-firms. Let the number of I-firms audited and then randomly chosen for cross-matching be \( a = As < m \), where \( m \) is now the number of I-firms in the long run equilibrium with evasion, \( A \) is the number of firms audited and \( s \) is the fraction selected for (sorting and) cross-matching. Since all I-firms have an equal chance of being audited or selected for matching, the probability that \( k \) of the \( z \) I-firms from which an F-firm makes purchases are audited and selected for matching is given by:

\[
h(m,a,z,k) = \frac{\Gamma(z,k) \Gamma(m - z, a - k)}{\Gamma(m,a)}, \tag{A1}
\]

where, for example, \( \Gamma(z,k) = \frac{z!}{k!(z-k)!} \), is the number of combinations of \( k \) objects out of \( z \) objects. Equation (A1) is a hypergeometric distribution with parameters \( m, a \) and \( z \).\(^{20}\) The upper and lower limits of this distribution are respectively \( y = \min(a, z) \) and \( y = \max(0, a - (m - z)) \). The second term in \( y \) merely recognizes that, if \( z \) is large, then some of the \( z \) I-firms from whom purchases are made will always be audited. Since \( m \) and \( a \) are given constants in this Appendix, we write \( h(z,k) \) for the expression in (A1) without risk of confusion.

Given that \( k \) I-firms are audited, what is the distribution of the number of I-firms whose sales are detected? The fraction of sales to an F-firm by such an I-firm that is revealed is \( \Theta/z \). Since the probability of detection on audit of any I-firm is \( e \), the probability of detection of the sales of \( j \) I-firms has a binomial distribution \( b(e,k,j) = \Gamma(k,j)e^j(1-e)^{k-j}, \) for \( j = 0,1,2, ..., k \). Without confusion, this probability is denoted \( b(k,j) \). With a successful audit, the remaining \( (1-\Theta)/z \) of sales by the I-firm to the F-firm come to light.

From Table 3-1, for unaudited F-firms, the STA can only compare the total value of sales invoices to an F-firm from audited I-firms with the total purchases reported by the F-firm. Denote the minimum number of I-firms that need to be detected to establish under-reporting by the F-firm, given that \( k \) relevant I-firms are audited by \( J(z,k,\Theta) \). For a given value of \( \Theta \), \( J(z,k,\Theta) \) is clearly weakly increasing in \( z \) and weakly decreasing in \( k \) and \( \Theta \). So expected under-reporting that will be detected, when \( k \) relevant I-firms are audited is

\[
\sum_{j=J(z,k,\Theta)}^{k} b(k,j) \left[ \frac{k\Phi}{z} + \frac{(1 - \Theta)j}{z} \right]. \tag{A2}
\]

Consequently, if \( J(z) = \min_{k} J(z,k,\Theta) \) is the minimum number of I-firms that need to be audited to detect under-reporting (\( J(z) \) is the smallest integer that is not less than \( z\Theta \)), the expected amount of under-reporting that will be detected is given by

\[
\sum_{K=J(z)}^{\infty} \sum_{j=J(z,k,\Theta)}^{k} b(k,j) \left[ \frac{k\Phi}{z} + \frac{(1 - \Theta)j}{z} \right]. \tag{A3}
\]

\(^{20}\)In the usual urn analogy, the urn contains \( z \) black balls out of a total of \( m \) balls. The number of balls drawn without replacement from the urn is \( a \). The hypergeometric distribution describes the number of black balls out of the \( a \) balls drawn from the urn.
Consider next the case of an audited F-firm. From Table 3-1, expected under-reporting that will be detected when \( k \) relevant I-firms are audited and selected for matching is given by \( k[\Phi + e(1-\Phi) - \Theta] / z \) (noting that the mean of the relevant binomial distribution is \( ke \)).

We can now put the pieces together. The probability of not being audit is \((1-p)\), the probability of an unsuccessful audit is \(p(1-e)\) and the probability of a successful audit is \(pe\). If \( \Theta < \min[Y/z, \Phi + e(1-\Phi)] \), expected under-reporting detected is, therefore, given by:

\[
pe(1 - \Theta) + p(1 - e) \frac{\Phi + (1 - \Phi)e - \Theta}{z} + (1 - p)(\text{something}) \]

In (A4) the fact that the mean of the hypergeometric distribution is \(az/m\) has been used in the second term. If \( \Theta \geq \min[Y/z, \Phi + e(1-\Phi)] \), then cross-matching is irrelevant for either audited or unaudited firms and either the third term or the second term of (A4) (or both) drops out. Given its report, \( \Theta \), a risk neutral F-firm will seek to choose \( z \) to minimize the quantity in (A4). The following proposition can now be stated.

**Proposition A1:** If the expected amount of under-reporting by an F-firm that will be discovered by the STA is given by (A4) then the purchase of intermediate inputs by an F-firm from all \( m \) I-firms weakly (strictly) dominates purchasing from fewer I-firms for risk neutral (averse) F-firms.

**Proof of Proposition A1:** Since the first two term in (A4) are independent of \( z \), we need to show that \( \delta(z, \Phi) \geq \delta(m, \Phi) \) for all \( z < m \) where \( \delta(z, \Phi) \) is defined as

\[
\delta(z, \Phi) = \sum_{k=z(z)}^{Y} h(z, k) \sum_{j=1(k, \Phi)}^{k} b(k, j) \left[ \frac{\Phi}{z} + \frac{(1 - \Phi)j}{z} - \Theta \right]
\]

and \( \delta(m, \Phi) \) is as in (A5) except that, since purchases are made from all matched I-firms, the distribution of purchases discovered through cross-matching collapses to \( b(a, j) \). Note first that, since \( Y \leq a \), if \( \Theta \geq a/m \), then by purchasing from \( m \) I-firms, the F-firm can ensure that no under-reporting is detected. Consequently, the proof is complete if the claim is shown to hold for \( \Theta < a/m \).

We show, by induction on \( z-J(z) \), that \( \delta(z, \Phi) \geq \delta(z+1, \Phi) \) for \( \Theta < a/m \). Now either \( Y = z+1 \) with \( z+1 \) transactions or \( Y = a \) with both \( z \) and \( z+1 \) transactions. It suffices to consider the first case that is more stringent. Secondly, with \( z+1 \) transactions, the lower limits of the binomial distributions in (A1) associated with each \( k \) are at least as great as with \( z \) transactions.

By direct computation, the inequality can be shown to hold for \( J(z) = z \), noting that, \( \Theta \geq z/(z+1) \) in this case. Now assume the inequality is true for \( J(z) = J' \).
The (cumulative) probability of at most \( k \) successes out of \( z \) possible successes with a hypergeometric distribution exceeds that of at most \( k \) successes out of \( z+1 \) possible successes for \( k \leq z \) since

\[
\text{h}(z,k) - \text{h}(z+1,k) = \text{h}(z+1,k) \left[ \frac{(z+1)\alpha - (m+1)k}{(z+1)(m-z-a+k)} \right]. \tag{A6}
\]

That is, the term in square brackets is positive at \( k=0 \) but decreasing in \( k \) and cumulative properties must sum to unity so that the graphs of the two cumulative distributions cannot intersect at \( k \leq z \). Consequently, the probability of at least \( k \) successes out of \( z \) possible successes is less than that out of \( z+1 \) successes.

Expanding \( \delta(.) \), we have, by assumption:

\[
\sum_{k=0}^{z} h(z,k) \sum_{j=1}^{\text{J}(z,k)} b(k,j) \left[ \frac{k\Phi}{z} + \frac{(1-\Phi)j}{z} - \Theta \right] \geq
\]

\[
\sum_{k=0}^{z} h(z+1,k) \sum_{j=1}^{\text{J}(z+1,k)} b(k,j) \left[ \frac{k\Phi}{z+1} + \frac{(1-\Phi)j}{z+1} - \Theta \right]. \tag{A7}
\]

Given (i) the property of the hypergeometric distribution discussed above and (ii) that the expected discovery given that \( z+1 \) relevant firms are audited and matched exceeds that when fewer firms are audited and matched: the term involving \( h(z+1,z+1) \) can be apportioned among the remaining terms on the right hand side so that A3 can be rewritten as

\[
\sum_{k=0}^{z} h(z,k) \sum_{j=1}^{\text{J}(z,k)} b(k,j) \left[ \frac{k\Phi}{z} + \frac{(1-\Phi)j}{z} - \Theta \right] \geq
\]

\[
\sum_{k=0}^{z} h(z,k) \sum_{j=1}^{\text{J}(z+1,k)} b(k,j) \left[ \frac{k\Phi}{z+1} + \frac{(1-\Phi)j}{z+1} - \Theta \right] + \text{Remainder}. \tag{A8}
\]

By the arguments above, the remainder is positive and decreasing in \( \Theta \). Clearly, therefore, A8 will continue to hold if \( \Theta \) is replaced by any smaller non-negative fraction, in particular values of \( \Theta \), say \( \Theta^* \), for which \( \text{J}(z)=\text{J}^*-1 \).

It only remains, therefore, to add terms for the case where \( \text{J}^*-1 \) firms are audited and matched to both sides of A3, noting that \( \Theta=\Theta^* \). Since \( \Theta<\alpha/m \), \( \text{J}^* \) cannot exceed the smallest integer at least as great as \( za/m \) so that \( \text{J}^*-1 < za/m \). Using A6 and this fact it is easily computed that the term added to the left hand side exceeds that added to the right. Thus, the desired inequality is true for \( \text{J}^*-1 \) firms if it is true for \( \text{J}^* \) firms. This completes the proof for risk neutral firms.

For risk averse firms note that if the expected discovery of under-reporting when purchases are made from \( m \) firms is identical to that when purchases are made from fewer than \( m \) firms, then the former will be preferred. This is because the distribution of discovered under-reporting with fewer than \( m \) firms can be obtained as a mean preserving spread of the distribution with \( m \) firms. In fact, the mean discovery of under-reporting with \( m \) firms is lower than that with fewer firms, strengthening the dominance of purchases from \( m \) firms. The weak
dominance result for risk-neutral firms is somewhat stronger than apparent from the proposition. When \( m \) is not too small, there will exist a number, \( M \), such that purchases from at least \( M \) firms strictly dominates purchases from fewer firms: Weak dominance is "almost strict". In view of this we assume, from section IV onward, that F-firms purchase intermediate goods from all I-firms. With this assumption, the distribution in the third term of (A4) collapses to a binomial distribution \( b(a,e,j) \). That is, with \( z=m \), the distribution \( h \) in (A1) collapses (trivially) into a discrete distribution: \( \text{prob}[k=a]=1; \text{prob}[k]=0 \) for \( k \neq a \). This simplifies (A4), since all terms with \( k \) not equal \( a \) vanish leaving, only the \( k=a \) terms. The distribution of I-firm sales detected through cross-matching the VAT is similarly a binomial distribution where the number of independent trials corresponds to the number of audited F-firms selected for cross-matching.\(^{21}\)

There are two important implications of the analysis here. Firstly, the dominance of purchases from all I-firms may no longer be true if there are economies of scale to be reaped from bulk purchases. Even so, our results establish that cross-matching may cause purchases to be smaller than optimal given evasion gains thus leading to lost scale economies. Second, since multiple purchases will only be feasible if the number of I-firms is large, the loss in surplus due to the sacrifice in purchase economies will be less severe in industries with few I-firms and absent in a monopoly, an effect running opposite to the usual pattern of deadweight losses due to suboptimal output decisions.

\(^{21}\) Thanks are due to John McLaren for helping us improve the exposition here.
Chapter 4  
Fiscal Pathologies and Federal Politics: Understanding Tax Arrears in Russia’s Regions

Daniel Treisman

I. Introduction

Why are governments better able to collect taxes in some parts of a country than in others? Tax collection is an essential function of any state, and inability to collect is synonymous with state crisis. Yet, in many countries in the developing—and even some in the developed—world, the success of collection varies geographically. In some regions, tax-payer compliance is impressively high, while in others it is disconcertingly low.

This issue arises with particular force in federal states, where the budgetary system—often taxation rights and sometimes collection agents—are divided between two or more levels of government. In such a setting, tax collection is not just a game between individual enterprises and the central tax agency. Especially in developing world federations in which law enforcement is imperfect, there is often a third player in the game—regional government. Both regional governments and enterprises share an interest in preventing revenues from leaving the region and in free-riding on federal provision of public goods financed by tax-payers in other regions. Understanding this game is of obvious importance for explaining the fiscal systems and performance of a range of countries, from Argentina and Brazil to India.

This chapter examines the question of regional variation in tax collection in one particularly vivid setting—the Russian Federation. I focus on one aspect of collection: the level of regional tax arrears, which includes both non-payment of taxes in defiance of the law and authorized deferrals. I am not concerned here with evasion per se, in which the tax obligation itself is concealed. Russia is a fruitful setting in which to study regional variation in arrears for several reasons besides its federal structure. First, the fiscal system—along with the economic

---

1 I gratefully acknowledge the support of the Smith Richardson Foundation, UCLA Academic Senate, and UCLA Center for European and Russian Studies.

1 While it would be preferable to separate the two, data do not permit. Authorized deferrals represent a falling, though still significant, share of the total. The lobbying capacity necessary to negotiate deferrals is likely to be related to the political resources necessary to defy central tax collectors outright, so studying the two phenomena together may not be too problematic.
and political systems—has been in flux since the early 1990s. During such periods of change, the underlying mechanisms that connect taxation and politics are unusually exposed. Second, the country is regionally heterogeneous in ways that facilitate interregional comparisons. Third, Russia’s fiscal crisis of the 1990s, many aspects of which lingered into 2000, was and remains one of the principal obstacles to economic recovery in that country.

Regional levels of tax arrears might vary for a number of reasons. These can be classified as factors that affect: (a) the enforcement capabilities and resolve of central government, (b) the ability of enterprises to resist, and (c) the willingness and ability of regional governments to collude with enterprises against federal collectors. In previous work, I emphasized the third set of factors. Treisman (2000a) modeled the interaction between enterprises, regional, and central governments in a context of imperfect law enforcement. In return for bribes, regional governments offer to protect enterprises against central tax collectors—by lobbying to forestall bankruptcy actions, coopting or corrupting local judges and locally based enforcement agents, and threatening the central government with regional strikes and demonstrations if local enterprises are harmed. In other words, regional governments sell “fiscal protection” to firms that locate in their regions, helping them to run up arrears with impunity. A number of predictions of this model were closely supported by observed outcomes in Russia. For instance, the rate of growth of regional tax arrears should increase with political decentralization but decrease with an increase in the region’s share of tax revenues. Underground economic activity should increase with fiscal decentralization but decrease with political decentralization. The rates of increase of arrears and estimated underground output during the 1990s in Russia did change in relation to changes in fiscal and political decentralization in the ways predicted.

This chapter looks specifically for evidence that such enterprise-regional government collusion can help explain the regional pattern of tax arrears, even when other factors related to central enforcement capacities and enterprise characteristics are taken into account. It finds some supportive evidence. The next section reviews very briefly the fiscal system and recent fiscal performance in Russia, demonstrating the interregional variation in tax arrears. Section III develops hypotheses about possible causes of such variation. Section IV tests these hypotheses, using a cross-section, time-series dataset on arrears in Russia between 1994 and 1997. Section V concludes.

II. Tax Collection in Russia in the 1990s

Tax collections fell quite dramatically in Russia during the 1990s. The absolute figures differ depending on how various revenues are classified, and they change over time as the GDP statistics are updated and corrected. But the general trends remain clear and unchanging. Consolidated budget tax revenues dropped from about 30 to less than 21 percent of GDP between 1992 and 1998 (see Table 4-1).

This was entirely due to a collapse of revenues at the federal level. Regional (including local) budget tax revenues started at 11.9 percent of GDP in 1992, and, after various ups and downs, ended 1998 at 11.5 percent of GDP. By contrast federal tax revenues fell from 17.9

---

percent of GDP in 1992 to 9.3 percent in 1998. (These figures classify as tax revenues receipts from foreign economic activity—foreign exchange tax, imports and export taxes—that were officially classified as “off-budget fund” revenues in 1992-3, and as “non-tax” revenues from 1995 on. If such revenues were excluded, the fall in federal tax revenues would be from 14.5 percent of GDP in 1992 to 8.8 percent in 1998.)

As is evident from these figures, the degree of decentralization of tax revenues increased dramatically during the decade. The regional share in total tax revenues first grew from about 40 percent in 1992 to around 55 percent in 1993-4; it then fell back to 49 percent in 1995, as the central government tried to recentralize in the context of macroeconomic stabilization; but it rose again subsequently to reach about 55 percent as of 1998.

Though the resulting level of consolidated budget tax revenues—at 20 percent of GDP—is not unusually low for countries at Russia’s level of economic development (and the figure would be almost 27 percent if non-tax revenues were included, and higher still if one added extrabudgetary funds), the federal budget is relatively small and the sharp fall led to a series of crises in the financing of central government. Funding for various federal programs was often sequestered in mid-year. The federal government was blamed for long delays in the wages of public employees in education, healthcare, and other areas, though research suggests that these may be better attributed to the strategic inflation of regional public employment by regional governments (Gimpelson et al. 2000).

Table 4-1: Tax Revenue in Russia, 1992-98 (% GDP)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Consolidated budget</td>
<td>29.8</td>
<td>25.5</td>
<td>26.4</td>
<td>25.0</td>
<td>22.8</td>
<td>23.0</td>
<td>20.7</td>
</tr>
<tr>
<td>Federal budget</td>
<td>17.9</td>
<td>11.5</td>
<td>11.9</td>
<td>12.7</td>
<td>11.0</td>
<td>10.0</td>
<td>9.3</td>
</tr>
<tr>
<td>Regional budgets</td>
<td>11.9</td>
<td>14.0</td>
<td>14.5</td>
<td>12.3</td>
<td>11.9</td>
<td>13.0</td>
<td>11.5</td>
</tr>
<tr>
<td>Memo: GDP, trillion Rs</td>
<td>19.0</td>
<td>171.5</td>
<td>610.7</td>
<td>1540.5</td>
<td>2145.7</td>
<td>2521.9</td>
<td>2684.5</td>
</tr>
</tbody>
</table>

Sources: Fiscal data are from working data of World Bank (April 1997) updated with figures from Freinkman et al. (1999), and from Ministry of Finance reports on budget execution 1997 and 1998. GDP data are from Goskomstat, Rossiiskiy Statisticheskiy Yezhegodnik 1999. Figures have been adjusted to make years comparable despite change in classification of foreign trade revenues: in 1992-3 off-budget fund foreign economic activity tax revenues (forex tax, import and export taxes) have been included; in 1995-6, “non-tax” revenues from foreign economic activity have also been included.
Much of the decline in tax collections can be attributed to an increase in tax arrears. As Table 4-2 shows, the total debt to consolidated budget grew from 1.8 percent of GDP in 1993 to 9.6 percent in 1998. While a significant part of the total represented deferrals of tax in the early years, from 1996 on the vast majority of debt represents to non-payments and fines for non-payment, mostly of VAT and profit tax.

A large part of tax revenues of both the federal and regional budgets come from taxes that are shared—most notably, VAT and profit tax. Shares in these accounted for 61.4 percent of total federal tax revenues in 1996, and 41.7 percent of total regional tax revenues (see Table 4-3).

<table>
<thead>
<tr>
<th>Table 4-2: Tax Arrears in Russia (end year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
</tr>
<tr>
<td>of which non-payments</td>
</tr>
<tr>
<td>on VAT</td>
</tr>
<tr>
<td>on profit tax</td>
</tr>
<tr>
<td>on excises</td>
</tr>
<tr>
<td>of which deferrals</td>
</tr>
<tr>
<td>Federal budget share of total tax debt (%)</td>
</tr>
<tr>
<td>Federal budget share of total tax revs. (%)</td>
</tr>
<tr>
<td>Regional tax debt as percent of gross regional product</td>
</tr>
<tr>
<td>maximum</td>
</tr>
<tr>
<td>minimum</td>
</tr>
<tr>
<td>mean</td>
</tr>
<tr>
<td>coefficient of variation</td>
</tr>
<tr>
<td>Memo: GDP</td>
</tr>
</tbody>
</table>


The sharing rates are set in annual budget laws. In 1994-97, the regional budgets were entitled to 25 percent of VAT and between 63 and 66 percent of profit tax revenues. Treisman (1999) and Shleifer and Treisman (2000) argue that the division of property rights over these taxes reduces incentives for the two levels of government to prioritize collection of them, and creates incentives for regional governments to collude with enterprises to deprive the federal government of its share. By law, incoming revenues from shared taxes are supposed to be divided between the budgets in proportion to the official sharing rates. The tax collectors and Treasury system are both officially under federal jurisdiction. Yet the actual federal receipts from

---

3 More precisely, there are separate federal and regional profit taxes levied on the same base, and collected by the same agency, the Ministry of Taxes and Collections (formerly the State Tax Service). Regional governments can lower the rate on its profit tax below a federally set maximum. In practice, this system works like one of tax-sharing.
these taxes have often fallen short of the mandated share, while regional receipts have exceeded the regional share. As Table 4-2 shows, the federal budget’s share of total tax debt has always been higher than its share of tax revenues, suggesting that incoming revenues were not always divided in the mandated proportions.

Table 4-3: Sources of Tax Revenue, 1996

<table>
<thead>
<tr>
<th></th>
<th>% of total tax revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Federal</td>
</tr>
<tr>
<td>VAT</td>
<td>46.5</td>
</tr>
<tr>
<td>Profit Tax</td>
<td>14.9</td>
</tr>
<tr>
<td>Excises</td>
<td>20.7</td>
</tr>
<tr>
<td>Personal Income Tax</td>
<td>2.3</td>
</tr>
<tr>
<td>Export and Import Duties</td>
<td>10.4</td>
</tr>
<tr>
<td>Corporate Property Tax</td>
<td>0</td>
</tr>
<tr>
<td>Other Taxes</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Source: Working data of World Bank. Total tax revenues here do not include "non-tax" revenues from foreign economic activity.

The rate of growth of arrears as a share of tax actually collected is also generally much higher for the federal budget than the rate for regional budgets. In the years 1993-98, the growth rate for federal arrears outstripped that for regional arrears in all years (see Table 4-4). In 1998, tax debt to the federal budget increased by almost 23 percent of that year’s actual federal tax receipts. Tax debt to regional budgets that year grew by less than 7 percent of receipts. (In part, this might be due to the accumulation of interest charges on unpaid taxes; if debt to the federal budget accumulated earlier, the growth path of such fines and penalties would be more significant than for regional budgets.)

Not only did arrears apparently accrue faster to the federal than to regional budgets, the tax service’s efforts to force payment from delinquent enterprises seemed more effective when the debt being collected was owed to the regional budgets. The rate at which enterprises voluntarily paid tax arrears on being presented with a demand from the tax service was about the same for regional and federal arrears. But when one compares the rates at which the service recovered arrears through the seizure of cash and other assets, recovery was more effective for regional arrears than for federal arrears in 1996 and early 1997, the period for which I had data (see Table 4-5).

---

4 As the OECD's 2000 Economic Survey: Russian Federation argues, regional governments avoided sharing in part by conducting budgetary operations in the form of debt offsets.
Table 4-4: Rate of growth of tax arrears

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth in tax debt as % of GDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i.e. (debt_t - debt_{t-1})/gdpt</td>
<td>1.69</td>
<td>1.98</td>
<td>2.71</td>
<td>3.33</td>
<td>2.13</td>
<td>2.88</td>
</tr>
<tr>
<td>of which growth in debt to federal budget (% points of GDP)</td>
<td>0.97</td>
<td>1.23</td>
<td>1.44</td>
<td>1.82</td>
<td>1.31</td>
<td>2.10</td>
</tr>
<tr>
<td>of which growth in debt to regional budgets (% points of GDP)</td>
<td>0.72</td>
<td>0.74</td>
<td>1.27</td>
<td>1.51</td>
<td>0.81</td>
<td>0.77</td>
</tr>
<tr>
<td>Growth in tax debt as % of tax revenues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth in tax debt to federal budget as % of federal tax revenues</td>
<td>6.6</td>
<td>7.5</td>
<td>10.8</td>
<td>14.5</td>
<td>9.2</td>
<td>13.8</td>
</tr>
<tr>
<td>Growth in tax debt to regional budgets as % of regional tax revenues</td>
<td>8.5</td>
<td>10.4</td>
<td>11.3</td>
<td>16.5</td>
<td>13.1</td>
<td>22.6</td>
</tr>
<tr>
<td></td>
<td>5.2</td>
<td>5.1</td>
<td>10.3</td>
<td>12.7</td>
<td>6.3</td>
<td>6.7</td>
</tr>
</tbody>
</table>

*January-December.*

Sources: Russian Economic Trends Database for 1993; Goskomstat Rossii, Rossiisky Statistichesky Yezhегодник 1999 for other years; tax revenues are those in Table 1.

Table 4-5: Reports of State Tax Service on Recovering Arrears (recovered from accounts of tax-payers and their debtors, from seizures of cash and from sales of property) and of Voluntary Payments by Tax Debtors

<table>
<thead>
<tr>
<th></th>
<th>% of arrears outstanding as of 1.1.96 recovered by STS by 12.31.96</th>
<th>% of arrears outstanding as of 1.1.97 recovered by STS by 5.1.97</th>
<th>% of arrears outstanding as of 1.1.97 voluntarily paid by debtor after receiving demand, by 5.1.97</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>57.9</td>
<td>5.4</td>
<td>22.9</td>
</tr>
<tr>
<td>Federal</td>
<td>49.8</td>
<td>4.7</td>
<td>23.0</td>
</tr>
<tr>
<td>Regional</td>
<td>68.0</td>
<td>6.2</td>
<td>22.7</td>
</tr>
</tbody>
</table>

Source: Data from A. Lavrov and STS May 1, 1997 report on arrears; Russian Economic Trends, 1997:1.

While the trend was upward in almost all of Russia’s 89 regions, arrears to the consolidated budget accumulated faster in some than in others. For instance, in 1997 total tax debt ranged from just 2.7 percent of gross regional product in Moscow city to 18.8 percent in Magadan Oblast. Though the range remained roughly similar during 1994-97, the years for which data were available, the coefficient of variation—a measure of dispersion—actually dropped quite sharply, suggesting an increasing clustering around the mean. Figure 1 shows the change in regional tax debt between 1994 and 1997. Those observations above the dotted line represent increases in debt as a share of gross regional product. Only Tyumen Oblast, home to the country’s oil and gas industries (and presumably the two autonomous oblasts within it), saw an actual decline in arrears during this period, as pressures by the central government on the leading oil and gas companies took their toll.
III. Explaining the Geography of Tax Arrears: Theory

Why might tax arrears accrue faster in some regions than in others? Causes for regional variation can be loosely classified according to whether they affect the center’s ability to enforce, enterprises’ ability to resist, or regional governments’ incentives to collude with enterprises against the central government. The division is somewhat arbitrary since the three actors are players in the same game and the enterprises’ ability to resist obviously depends on the center’s ability to enforce, and so on. It is offered merely as a classifying device.
Figure 4-1: Regional Tax Arrears, Russia 1994-

Tax arrears 1994 (% of gross regional product)

Tax arrears 1997 (% of gross regional product)
Central Enforcement

Economic Development

A common assumption in the literature on taxation is that collection of taxes is easier in more economically developed countries or regions (see for instance Edwards and Tabellini 1991). Superior record-keeping is to be expected in more modern sectors. Agricultural communities often have sizeable, non-commercialized production which is hard to monitor and tax efficiently. While this may tend to reduce the ability to assess taxes, it may also weaken enforcement once tax obligations have been set. I therefore looked for a relationship between tax arrears and both per capita gross regional product (lagged one year to reduce endogeneity problems) and the share of agriculture in total output (as of 1993).

Particular Sectors

Various scholars have argued that in Russia in the 1990s tax arrears on the part of enterprises in the fuel and energy sector were tolerated by the federal government as an implicit payoff for public services this sector provided (Treisman 1998, 2000b, Gaddy and Ickes 1998). In particular, the energy sector continued to provide fuel to politically sensitive rustbelt enterprises, army bases, hospitals and power generators even when these recipients did not pay their bills. For instance, while the net debt accumulated by insolvent clients to the electroenergy sector increased by 8.7 trillion December 1995 rubles between 1993 and 1998, the tax debt of the electroenergy sector grew by 7.6 trillion rubles. Net overdue trade credit extended by the gas sector increased during the same period by 3.8 trillion December 1995 rubles; its tax debt grew by 3.4 trillion rubles (Treisman 2000b). If, as argued, firms in the energy sector were being repaid via their tax debt for channeling implicit subsidies to politically important recipients, they would not need to fear serious government sanctions. I test for this by including variables for the shares of fuel and electricity sector enterprises in the regions' industrial output.

Distance from Moscow

In a country that spans 11 time zones, sheer physical distance from the capital might be thought to influence the effectiveness of federal monitoring and enforcement. Distance might also reduce federal eagerness to take tough enforcement measures by increasing the credibility with which sanctioned regions could threaten secession or lesser assertions of autonomy. To check whether this influenced the size of regional tax arrears, I included a variable for the distance—in degrees longitude plus degrees latitude—of a region's capital from Moscow.

Enterprise Resistance

Financial Distress

Enterprises will, of course, be less likely to pay their taxes on time if they are physically unable to do so—if they just do not have the money. Enterprises that are only able to sell output at a loss may not be able to pay taxes in full. Of course, profit tax will not be due if the profit tax
base is actually profit (in Russia, various expenses, such as advertising, could not be written off, so the tax-base was probably larger than profit). But other taxes were levied on revenue or payroll. Another reason enterprises might be unable to pay taxes—even though profitable in theory—is if their clients do not pay their bills on time. Enterprises to which large debts have accumulated might for this reason lack the liquidity to pay their taxes. I test these hypotheses by using variables for the proportion of enterprises in a region that were loss-making and the total debt to enterprises in the region as a share of gross regional product. To reduce problems of endogeneity, I lagged both of these variables by one year.

**Size**

For three reasons, large—and especially very large—enterprises will have greater ability to resist paying their taxes in full. First, major conglomerates will have greater resources to lobby government officials at the center to defer taxes and show leniency. There are economies of scale in such lobbying—initial expenses are significant, but marginal cost afterward may be relatively low. A corner-store cannot afford to send its director to Moscow to lobby the head of the tax service or a Finance Ministry department head, and if it did the entrepreneur would not get through the door. But a Gazprom or a Lukoil is so integrated into the political elite that requesting tax favors is just a part of business as usual. The second reason why large enterprises can resist more effectively is the greater political cost of forcing them into bankruptcy. Firms whose workforce runs into the thousands can threaten serious strikes or demonstrations. The third reason has to do with the enterprise’s relations with regional government, and so is described in the next section.

Operationalizing these variables with data available proved difficult, and the indicators I chose are at best imperfect approximations. The first is the share of regional industrial output produced by firms occupying a “dominant position” in the market for a particular good. A firm is classified under anti-monopoly law as having a “dominant position” if it is responsible for 35 percent of total nationwide sales of a particular good and is engaged in industrial activity within Russia. I used data for 1997, the first year for which they were available. Clearly, “dominant” firms will without exception be large, but many large firms will not be “dominant”. It is a highly noisy indicator. As a second attempt to get at this hypothesis, I included the share of workers in small enterprises as a percentage of total employment. I used the data for 1996 since the definition used by Goskomstat changed that year (as required by a 1995 law) and so previous data are not comparable. Small enterprises under the new definition are those with: up to 100 employees in industry, construction and transport; up to 60 employees in agriculture and science and technology; up to 30 in retail trade and consumer services; and up to 50 in wholesale trade and other sectors. This complicated definition is probably not what would have been chosen on theoretical grounds, but these are the data available.

**Collusion with the Regional Government**

A number of factors will affect how willing and able the regional governor is to collude with the enterprise, protecting it against federal tax collectors and bankruptcy courts.

---

Institutional Dependence of Governor

If the governor is more dependent on central political officials, he will be less likely to risk their displeasure by protecting local enterprises against the federal tax collectors. Governors will be more dependent if they are appointed by central officials rather than locally elected, and if they have predecessors who have been successfully fired by central order. The system for selection of Russia’s regional governors changed during the period analyzed. As of mid-1994, 27 of the country’s 89 regions had elected chief executives; by mid-1997, almost all did. As of the beginning of 1994, governors had been removed by central decree or pressure in 13 regions; by the end of 1996, the total had risen to 21 regions.6

By contrast, if the governor has an independent basis of legitimacy, such as ties to an ethnic movement, this may reduce the dangers inherent in conflict with the center. Similarly, the chief executive of a higher status region—an ethnic republic—may have less to fear from central administrative punishments. Within Russia’s asymmetric federalism, the 22 ethnic republics have greater autonomy and privileges than ordinary oblasts or krais. One might expect the leaders of republics, therefore, to use their institutional resources to defend “their” enterprises. To test these hypotheses, I constructed dummy variables for: whether the region was a republic, whether the governor had been elected (at least by the end of September of the given year; I judged that elections later than this would not give a new governor time to significantly affect tax payments), and whether a previous governor had been fired and forced out by central action by the end of the previous year.

Governor’s Political Affiliation

If the governor’s political views and electoral affiliations diverge sharply from those of the central incumbent, he may be more likely to defy that incumbent by helping to steal the federal budget’s resources. Major divergences of political views are likely to correlate with different networks or party support. Those elected with help from the communist opposition or support of communist voters would have little desire to increase federal revenues under the control of President Yeltsin. In fact, such regional leaders may feel it is in their political interest to stoke up public confrontations with the incumbent president. To test this hypothesis, I constructed a dummy variable for whether a communist governor had been elected that year (between December of previous year and November of this one—a communist elected in December would clearly not have time to affect tax collection in that year). I classified a governor as communist if he was elected with the support of the communists (KPRF or the opposition umbrella group, NPSR), if he was a member of the KPRF leadership, or if he had run on the KPRF’s list in a national election.

Data were compiled from McFaul and Petrov (1998), McFaul, Petrov, and Ryabov (1999), and the Institute for East-West Studies Handbook of Regional Executives.
Fiscal Incentives of Governor and Transaction Costs

Whatever the political relationship between a given governor and the central government, governors are likely to be affected by fiscal incentives. As discussed in the previous section, a large part of the tax revenues collected in Russia is then divided up between federal and regional budgets. Clearly, governors will have a stronger incentive to defy the central tax collector and support enterprises that do not pay if the regional budget receives a relatively small share of the taxes collected. Where the regional government has a larger stake in tax revenue collected, the governor should be less likely to collude with enterprises to reduce their tax payments. In explaining change in the level of tax debt, it is change in the region’s share of tax revenues that matters rather than the absolute level. To test this hypothesis, I use a variable for the change in the region’s actual share in tax revenues the previous year. I lag the variable since governors cannot react to such fiscal incentives instantaneously—what matters is not the official division rates so much as the actual rates, and these can only be guessed from the previous year’s experience.

At the same time, the governor’s decision whether to help protect individual enterprises must depend on the cost of negotiating collusive deals. These transaction costs will be much higher if the same tax base is divided up between hundreds of smaller firms and the governor must negotiate covert deals with each of them than if it is concentrated in a few large enterprises. Thus, besides increasing an enterprise’s resources to fight off federal tax collectors by itself, size reduces the unit cost for a regional governor of colluding with that enterprise.

IV. Statistical Results

To evaluate the evidence for each of these hypotheses, I ran a series of panel regressions. The dependent variable was the region’s total tax debt as a share of gross regional product for the years 1994 to 1997, the period for which data were available. I calculated coefficients by OLS, but report “panel-corrected standard errors,” as recommended by Beck and Katz (1995), which are corrected for panel heteroskedasticity and contemporaneous correlation. To reduce problems of autocorrelation, I included a lagged term of the dependent variable, as recommended by Beck and Katz (1996), who argue that this is preferable to using generalized least squares or other available techniques. The results should therefore be interpreted as explaining change in tax debt rather than its absolute level. I also included dummies for year in the regressions, to separate general trends from patterns associated with particular regions.

Certain deficiencies of the data need to be acknowledged. The dependent variable—total tax debt as a share of gross regional product—includes, as noted already, both authorized deferrals and unauthorized non-payments. In addition, it includes both actual non-payments and the fines and interest charges accumulated on them. The latter may be very significant; according to the OECD, citing the Ministry of Finance, penalties and fines represented 65 percent of total tax debt to the federal budget by the end of 1997 (OECD 2000, p.89). Since such penalties and interest charges are levied on tax debt accumulated earlier, regional differences in the earlier

---

*Panel-corrected standard errors are more accurate than standard errors computed by the Parks FGLS method for data of the kind analyzed in this paper (see Beck and Katz 1995). In STATA, I used the option pcse.*
years would be amplified in later years. The problem is only slightly alleviated by including a lagged term of the dependent variable. At the same time, if penalties and fines are imposed selectively, this might be a way that politically influential companies reduce their obligations to the state with consequences that look exactly the opposite of the other strategy—simply not paying and using influence to prevent punishment. These data problems are serious, and only the unavailability of other data and the importance of the question convinced me of the usefulness of analyzing them. One should not be surprised if few results are significant.

Another data issue was the decision how to normalize tax debt figures. One might use debt as a percentage of tax actually paid. But this actually measures the ratio of two quantities that are both likely manipulated by enterprises and regional governments. The manipulations, in addition, may act as substitutes—enterprises that conceal much of their tax base, and so pay little tax, will not need to run up arrears in order to avoid payment. Rather than try to interpret such a complicated variable, I chose to use debt as a percentage of gross regional product, a rough measure of tax debt as a percentage of the tax base. Gross regional product is also subject to manipulation if enterprises conceal part of their output to avoid taxes. But the distortion may be slightly less.

A third issue that raises problems for the statistical analysis that were not easy to deal with is the inherently interactive nature of the problem. If growth rates of tax arrears are the outcome of a game between individual enterprises, regional governments, and federal tax collectors, then what one really needs to model and statistically test for are interaction effects of variables related to each of the players. To introduce such interaction effects into the given model, though, would introduce unmanageable complexity. I settled, as a first attempt, for an oversimplified and probably somewhat misspecified model rather than for a complex estimation with numerous interaction terms for which the data did not provide enough degrees of freedom.

Table 4-6 shows the results. Model 1 contains just the variables I have argued may correlate with the strength of central enforcement; model 2 adds in factors related to enterprises’ resistance capacities; and model 3 adds in some possible indicators of regional government-enterprise collusion. Again, the classification of these variables is no more than an organizing device, and the important column to focus on is column 3, in which all variables are included.

The first point to note is the very strong continuity between years in the size of tax debt. The previous year’s tax debt was very significantly associated with the current year’s, and the estimated coefficient was very close to one. Interestingly, arrears did not show an explosive trend, as might have been expected given the self-compounding nature of fines and interest payments. The coefficient on previous year debt in column 3 suggests that arrears rose by only an unexplained two percent per year. In addition, tax debt tended to grow in 1996 relative to 1995, but there was no significant direction in 1997 (at least, that was not explained by other variables).

As feared, few significant results emerged. Regions with a higher concentration of output in the fuel or the electricity sectors did not have significantly faster accumulation of tax arrears, although the electricity sector variable did have the right sign, and was close to significant at the .10 level. Physical distance from Moscow also had the right sign but was not significant. The
regressions did not find any evidence that regions with more loss-making enterprises or larger accumulations of debt to enterprises had faster growing tax arrears. Nor did the regressions find any evidence that tax arrears accumulated faster in republics, in regions with elected leaders, or in regions where no previous leaders had been forced out of office by central action. Because of the serious data problems involved in this exercise, I am reluctant to view this as reason to reject these hypotheses and prefer to remain agnostic on them.

\footnote{In this and subsequent statements of the results, the reader should guard against committing the well-known ecological fallacy. One can \textit{not} assume that lack of a demonstrated relationship between loss-making and tax arrears at the regional level implies lack of a relationship at the enterprise level. The regressions presented do not test whether loss-making enterprises tend to have faster accumulating tax arrears.}
Table 4-6: What Determines Regional Tax Arrears?
Dependent variable is regional tax debt as percent of gross regional product. OLS with panel-corrected standard errors

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Central Enforcement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per capita gdp, previous year</td>
<td>-.070** (.031)</td>
<td>-.072** (.031)</td>
<td>-.065** (.031)</td>
</tr>
<tr>
<td>agricultural share in output (1993)</td>
<td>-.037 (1.63)</td>
<td>.44 (1.75)</td>
<td>-.16 (1.76)</td>
</tr>
<tr>
<td>share of output in fuel and energy sector (1997)</td>
<td>-.007 (0.09)</td>
<td>-.005 (0.09)</td>
<td>-.007 (0.09)</td>
</tr>
<tr>
<td>share of output in electricity sector (1997)</td>
<td>.006 (.011)</td>
<td>.018 (.012)</td>
<td>.019 (.012)</td>
</tr>
<tr>
<td>Distance from Moscow (in degrees latitude + degrees longitude)</td>
<td>.001 (.004)</td>
<td>.005 (.004)</td>
<td>.006 (.004)</td>
</tr>
<tr>
<td><strong>Enterprise Resistance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of enterprises loss-making previous year</td>
<td>-.013 (.012)</td>
<td>-.010 (.012)</td>
<td></td>
</tr>
<tr>
<td>Debt to enterprises previous year</td>
<td>.014 (.012)</td>
<td>.014 (.012)</td>
<td></td>
</tr>
<tr>
<td>Percent of regional output produced by domestic enterprises with more than 35% of sales nationwide of at least one product</td>
<td>.013 (.009)</td>
<td>.018** (.009)</td>
<td></td>
</tr>
<tr>
<td>percentage of workforce employed in small enterprises (1996)</td>
<td>-.042 (.034)</td>
<td>-.037 (.034)</td>
<td></td>
</tr>
<tr>
<td><strong>Enterprise-Region Collusion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Republic</td>
<td>.094 (29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>governor elected</td>
<td>-.023 (26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a governor had been fired by center in this region</td>
<td></td>
<td>.023 (26)</td>
<td></td>
</tr>
<tr>
<td>a (new) Communist governor elected in region this year</td>
<td>.65* (38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in region's share of taxes previous year</td>
<td>-.042*** (.015)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>previous year arrears/grp</td>
<td>1.06*** (.05)</td>
<td>1.00*** (.06)</td>
<td>1.02*** (.06)</td>
</tr>
<tr>
<td>1996</td>
<td>1.03*** (.26)</td>
<td>1.19*** (.26)</td>
<td>.60* (.33)</td>
</tr>
<tr>
<td>1997</td>
<td>.11 (.33)</td>
<td>.51 (.40)</td>
<td>-.10 (.49)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.74*** (.47)</td>
<td>1.73** (.71)</td>
<td>1.89** (.73)</td>
</tr>
<tr>
<td>Wald chi-square</td>
<td>872</td>
<td>923</td>
<td>979</td>
</tr>
<tr>
<td>Pr &gt; chi-square</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>231</td>
<td>230</td>
<td>230</td>
</tr>
</tbody>
</table>

Notes: * p < .10; ** p < .05; *** p < .01. Omitted year dummy is 1995.

1 "Communist" if elected with support of KPRF or NPSR, a member of the KPRF leadership or ran on its list in past election.
However, a few variables were significant with the expected signs. First, less developed regions (as measured by previous year gross regional product per capita) had faster accumulating tax arrears. (If the 1994 level of gross regional product was used, to reduce problems of endogeneity even more, the estimated effect was even more significant and slightly larger. The level of agricultural output, a second indicator of lower development, was not significant.) This effect was quite robust to the inclusion of additional variables, as comparing columns 1-3 shows. But it was not particularly large. A one standard deviation higher real per capita GRP in 1997 was associated with a .33 percentage point lower level of tax debt as a share of GRP. That year, tax debt as a share of GRP ranged from 2.7 to 18.8 percent.

Second, even using very imperfect indicators, there was some significant evidence that the size structure of a region’s enterprises affects the accumulation of tax arrears. The larger the share of regional industrial output produced by “dominant firms”—those industrially active within Russia and selling 35 percent of the market in a given good—the faster tax debt accumulated. Judging by the coefficient in column 3, a region in which “dominant firms” accounted for 50 percent of industrial output would have tax debt in a given year .7 percentage points higher than a region where “dominant firms” produced just 10 percent. Again, this is not an overwhelmingly large effect, but given the data imperfections it may be surprising to see any at all. The variable for the share of regional employment in small enterprises had the expected negative sign, but was not significant.

Third, regions that had just elected a communist governor did tend to have larger increases in tax debt that year. This was marginally significant (at p < .10). Such regions tended to have increases in tax debt about .65 percentage points higher than in regions without such an electoral history. One interpretation of this is that communist outsiders, when elected, are more likely to collude with local enterprises to defy the central tax collectors. The effect seems to be short-lasting: a dummy simply for whether or not the region had a communist governor was not significant. Various observers have noted that governors who come to power in Russian regions as hardline communists—Stroyev in Orel, Starodubtsev in Tula, even Rutskoi in Kursk—are usually coopted relatively quickly into the system of rewards and the game of prickly cooperation with the central government.

Finally, the column 3 regression suggests that regions whose share in total tax revenues increased most the previous year tended to have a slower accumulation of tax arrears. This was highly significant. A 10 percent increase in a region’s share of tax revenue was followed by a .4 percentage point lower increase in tax arrears the next year. This would seem to fit the hypothesis that giving regions a larger stake in tax collections increases the effectiveness of collection. To take a couple of examples, Kemerovo Oblast experienced a large drop in its share of tax revenues from 74 percent in 1995 to 66 percent in 1996. Its tax debt rose by an amazing seven percent of GRP in 1997, the largest increase of any region that year. Udmurtia’s share rose dramatically, from 62 percent in 1995 to 80 percent in 1996. Its tax debt grew relatively slowly in 1997, increasing by only .6 percent of GRP. The relationship seems, as hypothesized, to be with the change rather than the absolute level of the region’s tax share. When just the previous
year's regional share is included in similar regressions, it is not significant. If both the level (lagged two years) and change are included, the change remains significant (at \( p < .04 \)), with a slightly lower coefficient (-.036 instead of -.042); the second lag of the regional tax share is not significant at all.

\[ V. \text{Discussion and Conclusion} \]

Tax arrears grew dramatically in the mid-to-late 1990s in Russia. They grew more rapidly at the federal than the regional levels, and accumulated at widely different rates in different regions. These mounting arrears, at a time of dwindling federal fiscal resources, prompted repeated storming campaigns by the federal government, which tried to brow-beat debtor enterprises into paying up.\(^9\) It also led to sequestering of budget payments and the chronic underfinancing of federal programs (at least relative to what had been approved in the budget).

The lack of reliable, appropriately detailed data makes analyzing the causes of the accumulation of tax arrears extremely difficult. Nevertheless, some tentative examination of the figures is possible. From panel regressions of regions' tax debt, normalized by gross regional product, several positive conclusions emerge. A couple of factors point to the greater difficulties of collecting taxes in certain types of regions. Less developed regions had a significantly faster accumulation of tax debt than their more developed counterparts, and this was not explained by poorer financial performance in the less developed regions, which was controlled for. Either the federal enforcers are more lenient toward tax debtors in less developed regions, or the lower level of development makes it easier for these debtors to resist. Second, some evidence suggests that very large firms with significant market power may be better at reducing the amount of tax they pay than smaller enterprises (though here we must be careful to avoid the ecological fallacy). Regions with a more significant presence of such market leader firms and a smaller share of small enterprises tended to accumulate tax debts at a faster rate.

There is a second interpretation of this result, however: in regions where major firms are responsible for a larger share of output, this simplifies the task for regional governors of negotiating collusive deals. Rather than colluding with hundreds or thousands of firms to get them deferrals or hold off federal enforcers in return for political benefits or bribes, the regional government can negotiate with just one or two massive conglomerates. Transaction costs of such deals will be lower.\(^10\) The finding therefore also fits well with a view in which collusion between enterprises and regional governments plays a key role in explaining mounting tax arrears.

Two other factors also suggest the importance of regional governors in explaining the growth of tax arrears. First, regions in which a communist governor had just been elected tended to experience a larger than usual jump in tax arrears. And, most significantly, the growth of arrears seemed to be sensitive to the stake that regional governments had in the tax revenues collected. Where the regional share of tax revenues had recently fallen, arrears tended to grow faster than in regions where the regional share had recently increased. This suggests that property

\(^9\) For discussion of these, see Shleifer and Treisman 2000.

\(^10\) Though if the large enterprises are multiregional, they may be able to exploit competition between regional governors for their business, thus reducing the regional governor's share of the rents (Treisman 2000a).
rights and the incentives they create may be a key element in understanding the effectiveness of tax collection in Russia's regions.

The growth of tax debt slowed after the financial crisis of August 1998, during which higher inflation eroded the value of outstanding debt.\textsuperscript{11} Rising world prices for oil and other commodities also led to increases in receipts from export duties and other taxes on exporters. As a result, federal tax collections improved quite substantially in 1999. But, if as this paper's analysis suggests, the previous accumulation of tax debt was rooted in part in a particular political and fiscal interaction between regional governors and central officials, a return to higher tax arrears remains possible and will remain so until the fiscal system and the game of federal politics are reformed to remove such incentives.

\textsuperscript{11} In real terms, total debt to the consolidated budget fell in 1998 by slightly more than it increased in 1997.
References


Chapter 5

The Political Economy of Center-State Fiscal Transfers in India*

M. Govinda Rao and Nirvikar Singh

Abstract. India's federal system is distinguished by tax and expenditure assignments that result in large vertical fiscal imbalances, and consequent transfers from the central government to the state governments. Several channels are used for these transfers: the Finance Commission, the Planning Commission, and central government ministries. The transfers include statutory tax sharing as well as various categorical and block grants. While predetermined formulas are used for some transfers, there is considerable discretion in allocating other classes of transfers. In this paper, we use panel data on center-state transfers to examine two broad classes of hypotheses: (i) the economic importance of the states, measured by the state domestic product, influences the level and the composition of per capita transfers to the states; (ii) the political importance of the states, measured by their importance in the ruling coalition, and by whether the ruling party at the central and state levels is the same, influences the level and the composition of per capita transfers to the states. We control for variables such as population and per capita income in the panel regressions, and for state fixed effects, though population can also be considered a measure of political importance. We do find evidence supporting the bargaining view of federalism, as manifested in the result that states with indications of greater bargaining power seem to receive larger per capita transfers.

* This is a substantially revised version of a paper presented at the Columbia University-World Bank Conference on Institutional Elements of Tax Design and Reform, February 18-19, 2000, at Columbia University. It also forms part of a larger, ongoing study of Indian federalism by the two authors. The first version of these ideas was presented at the National Institute for Public Finance and Policy (NIPFP), New Delhi in September 1998. We are indebted to Bhaskar Dutta, John McLaren, Dilip Mookherjee, Rohini Pande, Ratna Sahay, and Barry Weingast, for detailed comments and suggestions on the February draft that led to (what we hope is) the improvement in this revision. Michael Dooley, Assaf Razin, Carlos Vegh, and Donald Wittman made useful suggestions when this work was presented at the UCSC SCCIE Development and International Economics Workshop, May 26, 2000. They are all definitely blameless for the remaining shortcomings. The second author would like to gratefully acknowledge the hospitality and support of the Research School of Pacific and Asian Studies at the Australian National University (Summer 1997), the NIPFP (September 1998) and the Center for Research on Economic Development and Policy Reform at Stanford University (Fall 1998 and Fall 1999), where his stays as Visiting Fellow have provided congenial environments for some of the research reported here. We are also grateful for financial support from ANU and the UCSC Academic Senate.
I. Introduction

William Riker, one of the most influential writers on federalism, advances an unequivocally instrumental view of the institution. He dismisses “ideological” reasons in favor of federalism, and strongly argues for understanding federalism as “a constitutional bargain among politicians”, with the motives being “military and diplomatic defense or aggression” (Riker, 1975, pp. 113-114). One may not accept the dismissal of noninstrumental motives, and instrumental motives realistically should also include broader economic concerns. However, the characterization of “a constitutional bargain” provides our starting point for examining the political economy of center-state fiscal transfers in India. While India is not the result of an explicit voluntary combination of its constituent parts, and while it is more centralized than many other federal systems, the functioning of India’s federal system does reflect implicit and explicit bargaining between the center and the states which make up the Indian Union.

Center-state fiscal transfers are a particular, quantifiable expression of the complex relationship between the national and subnational governments in India. These transfers are governed by a complicated mix of constitutional assignments, institutional precedents, discretion and negotiation. The result is therefore often difficult to understand or interpret. Analysts and policy-makers in India often focus on minutiae of formulas that govern various components of center-state transfers, or bargain behind closed doors over discretionary transfers. While detailed analysis of transfer formulas has normative significance, our approach in this paper, and our goal, are radically different. We take a positive approach, and we attempt to abstract from institutional details in our empirical exercise.

Specifically, we seek to examine the hypothesis that the political and economic influence or importance of the states affect the transfers they receive from the center. We do this with pooled cross-section time-series regressions, controlling parsimoniously for other factors that would affect transfers. We consider a simple measure of economic importance, the overall economic size of the state, as measured by State Domestic Product. We also consider demographic size (population) as a political variable, since the size of a state (viewed in a simplifying assumption as an irreducible unit). We also consider two alternative, more explicit measures of political influence. We do find some evidence to support the hypothesis that states with greater political and economic influence or importance receive higher per capita transfers. This is broadly consistent with the view of federalism as “a constitutional bargain among politicians,” even though the Indian system is not the result of an explicit joining of its constituent parts.

Riker (1975, pp. 156-158) discusses three arguments: that federalism promotes democratic polity, that it promotes democracy by promoting an interest in subnational government, and that it maintains individual freedom, and tends to dismiss them as “absurd”.

70
The structure of the paper is as follows. Section II fleshes out our hypotheses a little more, and includes a sketch of some additional literature that complements and extends Riker's view of federalism. We also note the limitations of our empirical exercise with respect to the hypotheses that we are able to examine. Since many will not be familiar with the Indian institutional context, we next review, in Sections III-V, some of the key features of Indian fiscal federalism and the system of center-state transfers. Section III summarizes some aspects of tax and expenditure assignments in India. We describe the constitutional assignments, and the constitutional provisions for center-state transfers, the need for which is substantially driven by the vertical fiscal imbalance created by the assignments. We explain the limited role of borrowing by the states, which helps increase their reliance on transfers.

Sections IV and V examine in more detail the institutions governing various kinds of center-state transfers. In Section IV, we give an overview of the institutions and their role in attempting to offset vertical fiscal imbalances. We discuss basic issues in the design of transfers, describe the volume and composition of transfers in India, and briefly examine the equity consequences of the transfer system. We also discuss the possible consequences of the tax sharing arrangements for tax effort. In Section V, we examine in more detail the Finance Commission, the Planning Commission, and the role of the central government through its various ministries in making transfers. We particularly analyze the functioning of the Finance Commission, because of its constitutional status, and the importance of its tax-sharing rules.

In Section VI, we present our empirical framework and results. We use panel data on center-state transfers to examine two broad classes of hypotheses: (i) the economic importance of the states, measured by the state domestic product, influences the level and the composition of per capita transfers to the states; (ii) the political importance of the states, measured by their population, their importance in the ruling coalition, and by whether the ruling party at the central and state levels is the same, influences the level and the composition of per capita transfers to the states. We control for variables such as per capita income in the panel regressions, and for state fixed effects.

We do find evidence supporting the bargaining view of federalism. Overall, the regressions suggest that, even with a very simple specification, variables that we believe measure bargaining power of the states, as components of the Indian federal system, are important. This conclusion is based on the positive estimated effect of economic and demographic size of the states on some components of per capita intergovernmental transfers, and of the lagged effect of a match between the state and central ruling parties on another component of transfers. In some specifications, there is also a positive effect of the proportion of ruling party/coalition MPs on per capita statutory transfers, again with a lag. Even for categories of transfers that are ostensibly governed by formulas, the pattern of transfers suggests that political bargaining factors are at work. This result is not surprising to anyone broadly in sympathy with Riker's view of federalism, or the many other political economy models of federalism, sampled in Section II. Nor are our results inconsistent with casual empiricism or political commentary on Indian

---

2 Sections III-V of the paper draw on Rao and Singh (1998a-d).
federalism. However, we believe our work is the first systematic empirical treatment and positive
evidence for these perspectives with respect to India. Section VII concludes with a summary of
our results and a discussion of further empirical work that is required.

II. Political Economy Models of Federalism

We began this paper with Riker's uncompromising view of federalism as a constitutional
bargain. Modelling this idea in full is daunting, but one can indicate some of its implications
quite simply. Consider a simple Nash bargaining framework, where the constituents of the
federation are the bargainers, and joining in the federation represents the cooperative outcome.
The threat points are, naturally, the welfare levels achieved in the absence of federation. Suppose
for simplicity of exposition that there are only two constituent units, and that utility is
transferable, so that total welfare from cooperation is independent of the distribution of the gains
from doing so. Let $u_i$ be the utility (welfare) that constituent $i$ gets from joining the federation,
and $d_i$ the disagreement payoff, or threat point. Let $C$ be the total welfare of the federation. We
may also allow bargaining power to be different across units, and it is denoted by $\lambda_i$. In this case,
it is easy to show that the Generalized Nash Bargaining Solution is given by:

(1) \[ u_1 = \lambda_1(C - d_2) + \lambda_2d_1, \]
(2) \[ u_2 = \lambda_1d_2 + \lambda_2(C - d_1). \]

If we normalize so that $\lambda_1 + \lambda_2 = 1$, then the gains from federation for the two units are given by:

(3) \[ u_1 - d_1 = \lambda_1(C - d_1 - d_2), \]
(4) \[ u_2 - d_2 = \lambda_2(C - d_1 - d_2). \]

We may view the rules of the federation as the practical means by which the final
outcome in (3) and (4) is implemented. Intergovernmental transfers are only one aspect of this
allocation of rules. In particular, tax and expenditure assignments, location decisions with respect
to centrally controlled economic activity, and explicit and implicit subsidies all affect the final
distribution of the benefits of the federal bargain. Unfortunately, there is no way that we can
imagine implementing a comprehensive empirical model that would test the full impact of this
complex of initial conditions and policies. Neither can we estimate the total surplus. Instead, we
view the observed pattern of intergovernmental transfers as one channel for distributing the
surplus created by federation. We consider the economic and political variables in the regressions
as proxies for, or determinants of, the threat points, $d_i$, and the bargaining powers, $\lambda_i$. Note that
the Nash bargaining model can be viewed as a normative outcome, determined by a set of
axioms. However, it may also be thought of as an approximate prediction of the outcome of a
strategic bargaining game. The only normative dimension that enters our empirical work is the
consideration of equity objectives (though they could also be conceivably driven by instrumental
concerns). These have received the most attention in writing on center-state fiscal transfers in
India, and we allow for them by including per capita income of the subnational units in the
regressions.
The above formalism is, of course, too simple in several respects. It does not explicate the institutional structures and rules that govern the sharing of the surplus created by cooperation. Also, it neglects the fact that the structures laid down at the time of federation can not specify fully how the surplus will be shared as it is generated over time. Finally, in a point that follows from the second one, federation creates a central authority that has the power to affect the distribution of the surplus across the constituent units of the federation, as well as an interest in preserving its power and the rewards that come with it. The analysis of such behavior of a central authority is part of a much more general strand of political economy, that examines rent-seeking, interest group behavior, and so on. We briefly describe how some of these issues are tackled in the literature, and the relevance for our empirical approach, which will, in the end, test a combination of hypotheses, not just the pure Rikerian postulate.

One strand of literature examines the economic determinants of nation formation, though not necessarily in the context of explicit ex post federations. An early contribution by Buchanan and Faith (1987) examines secession threats as a limit on taxation of groups within a jurisdiction. This is quite close to the Riker concept. Taxation is used for providing public goods as well as for redistributing surplus from those out of power to those in power. Only those in power receive transfers. The possibility of secession reduces the overall tax rate and the level of redistribution from what they would be in its absence. However, this analysis does not look at disparities within the ruling coalition, and how they affect transfers. Alesina and Spolaore (1997) examine nation formation as determined by a tradeoff between scale economies and costs of heterogeneity. They also examine compensation schemes to preserve or alter the boundaries of nations, but find that their assumptions rule out such schemes in equilibrium. Bolton and Roland (1997) also consider the potential breakup of nations, with an explicit focus on conflicts over redistribution. Since they allow for regions with heterogeneous preferences, they find that (unlike in the case of Buchanan and Faith’s model) poorer regions that receive transfers might still want to secede. In other words, the overall pattern of benefits and costs of federation matter, not just one component of redistribution. Finally, Wameryd (1998) examines the endogenous formation of jurisdictions in a rent-seeking model, and explicitly compares federalism, with hierarchical rent-seeking, to a centralized structure. These papers are only illustrative: Bolton, Roland and Spolaore (1996) provide a survey of this literature.

An alternative, but related, branch of literature examines distribution and redistribution in the context of existing nations, without the threat of secession or breakup being considered. Again, bargaining perspectives are important in this genre of models. Inman and Rubinfeld (1997) provide a transactions cost analysis of the federal provision of public goods. Their particular focus is on the role of legislative structures in determining this allocation. Given a clear assignment of tasks, a level of representation, and legislative institutions, one can compare the economic efficiency of different combinations of these three institutional variables. Building on the work of Breton and Scott (1978) and Baron and Ferejohn (1989), they make this comparison based on an assessment of different types of transactions costs. They do not explicitly treat intergovernmental transfers in their analysis. Kletzer and Singh (1997, 2000) analyze a median voter model of a federation with taxation, representative government, and
intergovernmental transfers. In their model, the constituent units of the federation realize that transfers have to be financed by taxes, and so they care about net transfers. They show in an example how coalitions may form to determine the winners and losers from transfers, based on factors such as income and agenda-setting power.

The analysis of Dixit and Londregan (1998) is the most complete treatment of fiscal federalism in the context of distributive concerns. They also provide an excellent survey of some of the literature in this area. In the Dixit-Londregan model, voters can belong to groups. They care about their private consumption as well as having ideological positions. They allow for political parties, and different political power of groups. The parties determine policies, including ideological positions as well as taxes and transfers. The political power of groups is positively affected by a greater willingness to compromise ideology for private consumption, and greater demographic importance at pivotal points in the preference distribution. Groups with greater power, measured in this way, are therefore predicted to do better in a federal system.

The models of distributive politics in an ongoing federation, whether they consider secession possibilities or not, are mostly positive, in the sense that the central government, as well as constituent units and individual voters, maximizes its own welfare. This is a feature of Buchanan and Faith's model, as well as Dixit and Londregan’s paper. Kletzer and Singh do not explicitly model the central government’s preferences, but assume that its goal is to stay in power, and it therefore responds to the median voter. Inman and Rubinfeld, on the other hand, provide a mix of positive and normative analysis. The important point is that no actor, including the central government, is assumed to be automatically benevolent in its objectives. This is the perspective in our paper, though we do provide critical comments on the efficiency of India’s current institutions governing center-state transfers.

What is the import of the above models for our empirical analysis? We believe that basic factors such as economic and political size matter for the kinds of distributive issues that are tied to the formation and continued existence of the federation, as we outlined simply in equations (3) and (4). Here marginal political power, or being pivotal in the ruling coalition, are less significant than basic bargaining power. In our empirical work, we attempt to capture this power through the impacts of the economic and demographic importance of states on per capita transfers. However, even where secession is not an issue, perhaps being too costly, the framework of federal institutions and rules provides an arena in which bargaining over distribution takes place. In this case, explicitly political variables that we consider, such as the proportion of the ruling party that comes from a state, or whether a state’s ruling party matches the party in power at the center, come into play. Thus our empirical work gets at both kinds of issues, albeit in an approximate way. The Dixit-Londregan type of analysis suggests that ideally, we would like to have data on the pivotal nature of parliamentary constituencies, aggregated by state. This data is available in raw form, but constructing the appropriate series will require considerable effort.
III. Tax and Expenditure Assignments

Constitutional Assignments

The Indian Constitution, in its seventh schedule, assigns the powers and functions of the center and the states. The schedule specifies the exclusive powers of the center in the Union list; exclusive powers of the states in the State list; and those falling under the joint jurisdiction are placed in the Concurrent list. All residuary powers are assigned to the center. The nature of the assignments is typical of federal nations. The functions of the central government are those required to maintain macroeconomic stability, international trade and relations and those having implications for more than one state. The major subjects assigned to the states comprise public order, police, public health, agriculture, irrigation, land rights, fisheries and industries and minor minerals. Subjects like public health, agriculture and irrigation involve considerable governmental expenditures. The states, being closer to constituents, also assume a significant role for subjects in the concurrent list like education and transportation, social security and social insurance.

The assignment of tax powers, however, is based on the principle of separation, i.e., tax categories are exclusively assigned either to the center or to the states (Appendix, Table 5A-1). Most broad-based taxes have been assigned to the center, including taxes on income and wealth from non-agricultural sources, corporation tax, taxes on production (excluding those on alcoholic liquors) and customs duty. A long list of taxes is assigned to the states. However, only the tax on the sale and purchase of goods has been significant for state revenues. The center has also been assigned all residual powers, which implies that the taxes not mentioned in any of the lists automatically fall into its domain.

Constitutional Provisions for Center-State Transfers

The Constitution recognizes that its assignment of tax powers and expenditure functions would create imbalances between expenditure needs and abilities to raise revenue. The imbalances could be both vertical, among different levels of government, and horizontal, among different units within a sub-central level. Therefore, the constitution provides for the assignment of revenues (as contrasted to assignment of tax powers), sharing of the proceeds of certain centrally levied taxes with the states, and making grants to the states from the Consolidated Fund of India. The assignment of tax revenues includes those taxes levied by the center, but with the proceeds assigned entirely to the states (Articles 268 and 269 of the Constitution). In one case (taxes on sale and purchase in the course of inter-state trade), the states have been allowed to collect the tax subject to a ceiling rate specified by the center.

The Constitution also provides for the compulsory sharing of the net revenue from non-corporate income tax (Article 270), and optional sharing of the proceeds of Union excise duty (Article 272). The shares of the center and the states and their allocation among different states of both the taxes are to be determined by the Finance Commission (see Section III, below) appointed by the President of India every five years, or earlier as needed. In addition to tax
devolution, the Finance Commission is also required to recommend grants to the states in need of assistance under Article 275.

**Revenue and Expenditure Shares**

The actual role of the central and state governments in revenue raising is summarized in Table 5-1. In 1997-98, the states on average raised about 31 per cent of total revenues, but incurred about 57 per cent of total expenditures. The revenues derived from exclusive central taxes constitute about 24 per cent, those from exclusive state taxes 37 per cent, from shareable sources 28 per cent and the remaining 14 per cent consists of non-tax revenues. The major taxes levied exclusively by the center consist of customs duty (15 per cent of total tax revenue) and corporation tax (8 per cent). Among the state taxes, the revenue from sales tax constitutes almost 17 per cent.

While the expenditure shares of central and state governments suggest a fairly high degree of decentralization, states' control over expenditure policies is less than the figures indicate. About 15 per cent of states' expenditures was on central sector and centrally sponsored schemes, which are specific purpose transfer schemes administered by various central ministries. States' expenditures on these schemes have actually shown an increasing trend, from about 7 per cent of total expenditures in 1985-86 to 15 per cent in 1995-96.

Over the last decade, while the share of the states in raising revenues has remained constant, their expenditure share has shown a steady increase, particularly since 1991, by about five percentage points. This has occurred because fiscal reforms initiated in 1991 have led to a deceleration in the growth of central government expenditures, but not so much in central transfers to states. Consequently, state expenditures have continued to increase even as central government expenditures decelerated.

**State Borrowing**

Article 293 of the Constitution does allow the states to borrow from the market. However, it stipulates that when a state is indebted to the center, it has to obtain the center's permission for market borrowing. As all the state governments are indebted to the center, states have no discretion. In practice, the Planning Commission, in consultation with the Union Finance Ministry and the Reserve Bank of India (RBI), has determined the total quantity of states' borrowing, and allocated the shares of each state.

Although in each state the overall transactions in a year should match revenues and expenditures, there are variations in daily and monthly positions. The cash balance position, or the "ways and means" position of the states is maintained by the Central Accounts Section of the Reserve Bank of India. States can also take overdrafts up to the limits stipulated by the RBI, by agreement with the state government. Any borrowing beyond this limit is called an 'unauthorized overdraft'. Until 1985, the states could resort to this means rather liberally. Further, when the overdraft position reached very high levels, from time to time, the center simply cleared the overdrafts by converting them into medium term loans. In January, 1985, an overdraft regulation
scheme was introduced, which stipulated that if the states continued to have the overdrafts with the RBI for more than seven continuous working days, the RBI was not obliged to honor the checks of such states.\(^3\) This measure has vested the center with more effective control of state borrowing powers, and at the same time has introduced harder budget constraints for the states.

**IV. Intergovernmental Transfers: Overview**

Economic arguments for intergovernmental transfers have been made in terms of (a) offsetting fiscal imbalances or closing fiscal gaps; (b) establishing horizontal equity across the federation and (c) offsetting inter-jurisdictional cost and benefit spillovers. In addition, transfers may also be given to carry out some agency functions for the central government. We focus here on the first of these arguments, in the Indian context.

**Offsetting Fiscal Imbalances**

"Vertical fiscal imbalance" refers to the difference between expenditures and revenues at different levels of government. Vertical fiscal imbalances are a feature common to all multilevel governmental systems. Even when assignments of functional responsibilities and revenue powers are efficient, imbalances are bound to occur. This is because the efficient expenditure assignment does not typically match with the efficient tax assignment. The central government has a comparative advantage in raising revenues and monitoring intergovernmental competition to control "free-riding" whereas sub-central governments are better placed to provide public services efficiently corresponding to varying preferences of people of different jurisdictions (Breton, 1987, 1996). Therefore, assignments according to comparative advantage result in vertical fiscal imbalance. Of course, actual assignments are the result of many non-economic considerations and these can contribute to vertical fiscal imbalances as well.

Vertical fiscal imbalances can also be caused by factors other than assignments. At the sub-national level, intergovernmental tax competition can result in lower tax rates, but competition to provide public services can enhance expenditure levels, thereby accentuating vertical fiscal imbalance. In addition, variations in fiscal management in terms of tax effort and expenditure economy among different levels of government can also contribute to the degree of vertical fiscal imbalance. For example, the states have failed to exercise their option of imposing an agricultural income tax.

Although the concept is intuitively clear, numerical measurement of the coefficient of vertical fiscal imbalance is problematic, because it is difficult to judge the degree of independence of various sources of revenue. For example, some types of tax revenue collected by a state government might be subject to restrictions on the base and rates imposed by the higher level government. Here, we measure vertical fiscal imbalance in India by taking only the own

---

\(^3\) The states reacted to this measure by taking resort to short term borrowing from the private sector, or from their own enterprises. The West Bengal government, for example, took short term loans from Peerless Insurance Company, a private sector financial firm.
revenues in the current account as the states' independent revenue source. There was a clearly increasing trend in vertical fiscal imbalances: the ability of the states to finance their current expenditures from their own sources of revenues has shown a decline from 69 per cent in 1955-56 to around 55 per cent in the 1990s (Table 5-2). During this period, the states' shares of current expenditures as well as current revenues have remained more or less constant, but the measure of vertical balance shows a decline. This apparent paradox is due to the fact that an increasing proportion of expenditure of central and state governments over the years has been financed from borrowing. Thus, the declining share of the states' own revenues to their current expenditures shown in column 4 actually reflects an increasing tendency to divert capital receipts to meet current expenditures.

Increasing centralization and imbalance are not entirely reflected in the quantitative indicators. The states' control over expenditure decisions has also eroded because the proportion of specific purpose, matching transfers for Central Sector and Centrally Sponsored Schemes in total state expenditures has increased, from 12 per cent in the fifth Plan (1969-74) to 18 per cent in the seventh Plan (1985-90) and, more recently, almost 15 per cent. Such transfers also have a longer term effect, when transfers under the schemes are no longer available, as the states can get locked into these centrally initiated schemes.  

The filling the "fiscal gap" rationale for intergovernmental transfers extends to horizontal fiscal imbalances as well. Horizontal fiscal imbalances refer to the mismatch between revenues and expenditures of governmental units within a level of government. In the Indian context, they refer to an excess of expenditures over revenues of different state governments. From the national point of view, the persistence of large horizontal imbalances has been considered improper, and these imbalances have been sought to be corrected through equalizing transfers from the center, which automatically implies the existence of some degree of vertical imbalance as well.

Horizontal fiscal imbalances can arise due to revenue or expenditure differences between the states. Revenue differences can be due either to differences in fiscal capacity or in effort. Similarly, expenditure differences between states may be due to differences in the quantity or quality of public services provided or differences in the unit cost (either due to factors beyond the control of the states or differences in fiscal management). Therefore, horizontal imbalances are not exogenous to the states' fiscal management, and do not, by themselves, provide a rationale for intergovernmental transfers.

Table 5-3 presents differences in per capita incomes, revenues and expenditures as well as poverty levels among the 25 states in the Indian federation. To facilitate meaningful comparisons the states have been classified first in terms of 15 relatively homogeneous ones and 10 'Special Category' states (seven North-Eastern States, Sikkim, Jammu & Kashmir and Himachal Pradesh). The former are again classified into high income, middle income and low

---

4 However, as John McLaren has pointed out to us, central control of these transfers can be interpreted as disguised central expenditure, which is another way of reducing the vertical fiscal imbalance created by the constitutional assignments.
income categories. The differences in revenues and expenditures presented in Table 5-3 bring out a number of features. First, there are wide variations in revenues between different categories of states, as well as among the states within each of the categories. Second, these variations indicate both inter-state differences in revenue capacity and in efforts to raise revenues. Thus, the variation in per capita taxes is much higher than that of per capita SDP. Also, some of the richer states have revenue - SDP ratios lower than middle and low income states in spite of the fact that the richer states have an advantage in exporting the tax burden to poorer states (as argued in Rao and Singh, 1998a). Third, per capita expenditure variations among the general category states (if the small state of Goa is excluded) are lower than the variations in per capita SDP. Fourth, the tax-SDP ratios in the special category states are lower than in the general category states in spite of their higher per capita SDP, partly because their SDP is derived mainly from government administration, outside the tax base. Finally, generally, the fiscal dependence of the states on the center was not only high, but also varied inversely with per capita SDP.

Design of Transfers

As mentioned, the fiscal imbalance argument does not exclude considering the actual fiscal behavior of the states. Designing transfers to offset fiscal imbalances can adversely affect incentives for own-revenue raising and for spending control (Wilde, 1971, Gramlich, 1977), and these disincentive effects ideally should be minimized. General purpose transfers are given to enable the sub-national governments to offset the fiscal disadvantages arising from a lower revenue capacity and a higher unit cost of providing public services. This can be achieved by giving unconditional transfers in a variety of ways, but the least distorting way is to give transfers equivalent to the recipient’s “need-revenue” gap (Bradbury, et. al., 1984). The need-revenue gap measures the difference between what a state ought to spend to provide specified levels of public services and the revenue it can raise at a given standard level of tax effort. In addition to avoiding incentives for fiscal laxity, a formula for intergovernmental transfers should be equitable, simple, transparent, and perceived to be objective. However, such ideal transfer systems do not exist in practice. In the actual design of transfers, historical, political, cultural factors can play important roles, so that simple normative criteria, even if agreed upon, may not easily translate into transfer systems that achieve the objectives. Trying to pick up the effect of some of these factors is, of course, the goal of our empirical exercise.

Volume and Composition of Transfers to States

Transfers from the central government contribute significantly to state finances (Table 5-4). In per capita terms at constant (1981-82) prices, central transfers to the states increased by over 2.5 times, from Rs.77 in 1975-76 to Rs. 194 in 1993-94, declining marginally thereafter owing to fiscal compression. Also, until 1993-94, the growth of transfers was faster than both the center’s and the states’ own revenues. Thus, the share of transfers in central revenues increased from 32 per cent in 1970-71 to 44 per cent in 1993-94, and then declined to 39 per cent in 1995-96. Similarly, the share of transfers in state revenues increased from 39 per cent to 44 per cent.

---

5 There are valid reasons for excluding Goa from such comparisons. Besides its small size, until the late 1980s, it was a Union Territory and required substantial initial spending to reach the status of a state.
and declined to 38 per cent in 1995-96. State expenditures increased even faster during this period, so that the share of transfers in state expenditures declined steadily. However, they still finance almost a third of state expenditures.

A notable feature of India’s federal fiscal arrangements is the existence of multiple channels of transfers from the center to the states. First, as mentioned earlier, the Finance Commission decides on tax shares and makes grants. Second, the Planning Commission makes grants and loans for implementing development plans. Finally, various ministries give grants to their counterparts in the states for specified projects either wholly funded by the center (central sector projects) or requiring the states to share a proportion of the cost (centrally sponsored schemes).

Historically, as development planning gained emphasis, the Planning Commission became a major dispenser of funds to the states. As there is no specific provision in the constitution for plan transfers, the central government channeled them under the miscellaneous and ostensibly limited provisions of Article 282. Before 1969, plan transfers were project-based. Since then, the distribution has been done on the basis of a consensus formula decided by the National Development Council (NDC). However, various central ministries still felt the need to influence states’ outlays on selected items of expenditure through specific purpose transfers with or without varying matching requirements: these are monitored by the Planning Commission.

The relative shares of the three channels of central transfers to states since the Fourth Plan, presented in Table 5-5, bring out some important features. First, there has been an increase in the discretionary element of transfers. The proportion of transfers recommended by the Finance Commission (statutory transfers to total current transfers) declined from 65 per cent during the Fourth Plan (1969-74) to less than 60 per cent during the Eighth Plan (1991-95). Formula based transfers from the Finance and Planning Commissions together have also tended to decrease relative to discretionary transfers, going from about 85 per cent of total transfers in the Fifth Plan period to 78 per cent during the Seventh Plan period. Specific purpose transfers for central sector and centrally sponsored schemes, constitute the bulk of discretionary transfers. The share of these transfers increased steadily, from less than 12 per cent in the Fourth and Fifth Plan periods, to about 20 per cent in 1994-95. Most of these schemes require matching contributions from the states. Thus, there is clear evidence of an increase in the discretionary element in transfers to the states, as one of the most significant political economy features of the intergovernmental transfer system in India.

Second, within statutory transfers, the proportion of tax devolution, which had already been high, has shown a steady increase while that of grants has declined. Thus, even as the share of statutory transfers declined from about 67 per cent during the fifth plan to about 57 per cent in

---

6 The legitimacy of these transfers has been seriously questioned. Some constitutional experts argue that transferring funds to the states under Article 282 is unconstitutional. Others consider that though this is permissible, channelling large amounts under this article is not in keeping with the spirit of the Constitution (See, NIPFP, 1993).

7 The NDC is chaired by the Prime Minister and its members include all cabinet ministers at the center, Chief Ministers of the states, and members of the Planning Commission.
1993-94, the share of tax devolution increased from 50 per cent to 53 per cent. Tax devolution constituted 84 per cent of statutory transfers during the Fourth Plan, but increased to almost 90 per cent during the Eighth Plan. Much of the increase in real per capita transfers (at constant prices), shown in Table 5-1, was from the increase in tax devolution. This may be explained by the fact that, while the Finance Commissions since the Seventh attempted to impart greater progressivity in tax devolution, this was done subject to the constraint of protecting the transfers of the better-off states in absolute real terms. In the event, both tax devolution and overall per capita transfers showed a significant increase.

**Progressivity of Transfers**

While theoretical considerations require that intergovernmental transfers should be designed to offset revenue and cost disabilities fully, actual transfer systems fall far short of this ideal. In practice, historical and political factors are at least as important as economic considerations in determining the transfer system. Thus, the volume of transfers made, the form and composition of the transfers, and the degree of progressivity in their distribution are all determined as a compromise between economic considerations and the constraints imposed by non-economic factors. We briefly focus on the equity objective.

The correlation coefficients and the cross-section income elasticity of different types of transfers with per capita SDP, beginning with the Sixth Finance Commission, are presented in Table 5-6. Several insights emerge. First, per capita transfers are inversely related to per capita SDP from 1979-80 onward. The absolute value of the (negative) correlation coefficients, and their significance levels, however, are higher in more recent years. Second, Finance Commission transfers had the highest progressivity. In fact, it is only since the recommendations of the Seventh Commission in 1979-80, that Finance Commission transfers and, consequently, total transfers, have had a significant negative correlation with per capita SDP. This is explained by the increased weight given to the backwardness factor in tax devolution by Commissions since the Seventh, as discussed below.

**V. Intergovernmental Transfers: Institutions**

**Finance Commission Transfers**

So far, ten Finance Commissions have made recommendations and, barring a few exceptions, these have been accepted by the central government. However, the working of these Commissions, their design of the transfer system, and the approach and methodology adopted by them have come in for criticism. The main criticisms are: (i) those relating to attempts to restrict the scope of the Finance Commissions through the Presidential terms of reference; and (ii) those on the approach and methodology employed by the Commissions and the equity and incentive consequences of the transfer scheme evolved by them.

India’s adoption of a planned development strategy with a pronounced socialist bias concentrated economic power in the hands of the center and, within the central government, the
Planning Commission. The increased dominance of the Planning Commission in allocative decisions, and its empowerment to dispense assistance to the states to finance their developmental activities, curtailed the Finance Commission's role in making intergovernmental transfers. Although the Constitution makes no distinction between Plan and non-Plan sides of the budget, and puts transfers under Articles 270 (income tax), 272 (excise duty) and 275 (grants) entirely within the jurisdiction of the Finance Commissions, Presidential terms of reference have confined the Finance Commissions to making transfers only to meet the non-Plan requirements of the states. The conflict in the jurisdictions of the two Commissions surfaced for the first time when the Third Finance Commission made its recommendations. Although its majority recommended inclusion of 75 per cent of the plan requirements of the states, the central government rejected this recommendation and accepted the recommendation of the Member-Secretary to avoid the plan side of the states' fiscal requirements altogether. For the subsequent Commissions (until the Ninth) the terms of reference themselves excluded the plan side from their scope. Even when the terms of reference did not impose any restrictions, as in the case of Ninth Finance Commission, the convention of assessing the non-plan side separately from the plan side was continued by precedent.

The restriction of the Finance Commissions to the non-plan side of the budget has led to a number of problems. First, larger transfers through the Planning Commission have significantly reduced the ability of the Finance Commission to achieve redistribution for fiscal equity. Second, it has prevented a comprehensive periodic review of state finances. Third, conceptually, the plan and non-plan distinction is unsound. Besides poor co-ordination, the separate treatment of plan and non-plan expenditure needs, and the emphasis on having large plans have led to inadequate provision for, and maintenance of, assets created under previous plans. From the states' point of view, separate plan and non-plan assessments gave them the opportunity to submit different projections to the two Commissions - an overestimated non-plan budgetary gap to the Finance Commission and overestimated saving in the non-plan account to the Planning Commission.

The Finance Commissions' approach consists of: (i) assessing overall budgetary requirements of the center and states to determine the volume of resources available to the center for transfer, and required by individual states; (ii) projecting states' own revenues and non-plan current expenditures; (iii) distributing assigned taxes, broadly on the basis of origin, (iv) distributing sharable taxes; and (v) making up the deficit between projected expenditures and revenues after tax devolution with grants. This is popularly known as the "gap-filling" approach.

Assigned taxes are distributed according to the principle of origin and there are no serious problems associated with them, though the states contend that the center has not exhausted the potential of Article 269 taxes (levied by the center but with proceeds assigned to the states). As

---

8 The grants \( G_i \) receivable by the \( i^{th} \) state are given by, \( G_i = E_i - (R_{oi} + R_{si} + R_{ei}) \), if the right hand side is positive, where \( E_i \) denotes projected non-plan current expenditures of the \( i^{th} \) state; \( R_{oi} \) = Projected own revenues of the \( i^{th} \) state; \( R_{si} \) = Projected share of assigned revenues of the \( i^{th} \) state; and \( R_{ei} \) = Projected shared taxes of the \( i^{th} \) state.
regards shared taxes, the basic issue is that, as the center gives away large shares to states, it concentrates on non-sharable revenue sources which not only creates horizontal inequities and relative price distortions, but also distorts the tax structure.

In the evolution of the system of tax devolution over the years, some important features are notable. First, states have always preferred tax devolution to grants, due to its inherent responsiveness to price and income increases. Second, the Finance Commissions, in response to criticism that their transfers promoted laxity in the states’ fiscal management, have preferred to increase tax devolution rather than “gap-filling” grants. For these reasons, tax devolution has shown a significant increase both in absolute terms and in relation to grants, with the states’ share of income tax increasing to 87.5 per cent and that of Union excise duties increasing to 47.5 per cent under the recommendation of the Tenth Finance Commission. These shares are distributed to the states mainly on the basis of general economic indicators like population, per capita GDP, other indicators of backwardness, collection and tax effort. In the final step, Article 275 grants fill “fiscal gaps.”

The shareable portion of the two taxes, and the criteria for their distribution among the states adopted by the ten Finance Commissions, are summarized in Appendix Tables 5A-2 and 5A-3. As already mentioned, tax devolution was made on the basis of general economic indicators like population and backwardness and not on the basis of fiscal disadvantages per se. Until 1989-90, tax devolution was not linked to the fiscal needs of the states as measured by the Finance Commissions at all. The Ninth Finance Commission, however, felt the need to link tax devolution to estimated deficits, but assigned this factor only 5 per cent weight. The Tenth Finance Commission increased this weight to 7.5 per cent. Of course all the Commissions assigned weights to population, which broadly represents expenditure need, but in doing this, the Finance Commissions were mandated to use the 1971 population figures, to provide an incentive to those states succeeding in controlling population growth. In the process, states that had higher population growth due to immigration, and not just higher birth rates, were also penalized.

Assigning weights to contradictory factors like “collection” and “backwardness” in the same formula for distribution has rendered the achievement of the overall objective of transfers difficult. While the objective of basing transfers on general economic indicators was to keep the devolution package simple and transparent, the purpose was lost when the Finance Commissions used multiple variables with different exponential powers, as was done in the case of inverse and distance forms of per capita GDP. The “backwardness” criterion included five variables in the Fifth Finance Commission. The Ninth Commission’s second report used three overlapping variables: scheduled caste and scheduled tribe population, agricultural laborers, and people below the poverty line.

Grants recommended by the Finance Commissions (Article 275) are determined on the basis of projected gaps between non-plan current expenditures and post-tax devolution revenues. Some of the Commissions moderated the “gaps” by taking account of normative growth rates of revenues and expenditures in projections, and taking the returns from public undertakings on a normative basis. Some of the Commissions (particularly after the Sixth) also attempted to
enhance outlays on specified services in the states by making closed-ended, specific purpose non-matching grants. However, these attempts were selective and limited. The Ninth Finance Commission was the first to attempt and comprehensively adopt the normative approach and determine the gaps between revenue capacities and expenditure needs, but since the bulk of the transfers was given through tax devolution based on general economic indicators, the effectiveness of this approach was dissipated. The Tenth Finance Commission reverted to the old methodology, on the grounds that it was not mandated to follow the “normative” approach in its terms of reference.

The “gap-filling” approach suffers from a number of shortcomings. First, Finance Commissions have made judgments about tax shares without evolving objective criteria for evaluating the center’s needs. Second, as noted earlier, the separate workings of the Planning and Finance Commissions have prevented an integrated view of the states’ fiscal needs, and distorted their behavior. A third weakness of Finance Commission transfer schemes is their lack of clear purpose. They have not been designed to meet the major theoretical objective of unconditional transfers, offsetting fiscal disadvantages of the states. Tax devolution was decided on different considerations from those of grants-in-aid, and the criteria for distributing the income tax were different from those for excise duties. Fourth, although successive Commissions assigned higher weights to backwardness in the tax devolution formula, general economic indicators still dominated. The methodology has also had an inherent bias against poorer states, because projections of budgetary gaps were made based on existing revenues and non-plan expenditures (representing low levels of services in states with lower tax bases).

Fifth, the gap-filling approach adopted by the Finance Commissions has had adverse effects on incentives. The center has tended to concentrate on non-shareable sources of revenue like import duties, thereby distorting the pattern of resource allocation. Similarly, the center has tended to mobilize resources by increasing administered prices on public monopolies rather than increasing excise duties on them, altering relative prices in unintended ways. At the state level, the gap-filling approach is said to have led not only to disincentives for tax effort, but also profligacy in spending.

Recent Commissions have modified the criticized approach in several ways. First, they substantially enhanced tax devolution so that very few states were left with post-devolution gaps to be filled by grants, and they tried to target tax devolution to poorer states. However, since, an overwhelming proportion is given on the basis of general economic indicators, the overall effect has been quite limited. Second, recent Commissions introduced selective norms for the center and the states by targeting the rates of growth of revenues and expenditures, and by assuming certain rates of return on their loans and investments. However, in the absence of a mechanism to enforce them, these have been merely of academic interest. Finally, whether or not the devolution system has in fact distorted the tax system in India, it does provide a perverse incentive to the center in its revenue raising. The tenth Finance Commission recommended an

---

9 For arguments on these lines see, Burgess, Howes and Stern (1993). Dasgupta and Mookherjee (1994), however, provide evidence against this hypothesis.
alternative scheme of devolution by pooling revenue all central taxes and giving a fixed share to the states and the government is in the process of implementing this recommendation. The proposed new arrangement is better than the prevailing system, though the problem of disincentives would not be entirely eliminated. The center still would have the incentive to raise non-tax revenues as against tax revenues, and raise administered prices of public monopolies instead of raising excise duties.

Plan Transfers

Plan transfers from the center to the states consist of grants and loans. Since 1969, plan assistance has been distributed on the basis of the “Gadgil formula” approved by the National Development Council modified from time to time. The latest modification in the formula was made in December, 1991. According to this latest formula, at present 30 per cent of the funds available for distribution is kept apart for the special category states. Assistance to them is given on the basis of plan projects formulated by them and 90 per cent of the transfer is given as grants, with the remainder as loans. The 70 per cent of the funds available to the major states is distributed with 60 per cent weight assigned to population, 25 per cent to per capita SDP, 7.5 per cent to fiscal management and the remaining 7.5 per cent to special problems of states. Of the 25 per cent weight assigned to per capita SDP, the major portion of the funds, 20 per cent is allocated only to the states with less than average per capita SDP on the basis of the “inverse” formula; the remaining 5 per cent of the funds is assigned to all the states according to the “distance” formula. For the major states, assistance is given by way of grants and loans in the ratio of 30:70. The transfers given to the states for plan purposes are not related to the required size or composition of plan investments (see Appendix Table 5A-4).

The Planning Commission works out five-year-plan investments for each sector of the economy and each state. With this as background, the states work out their respective annual plans for each year, based on the estimated resource availability, which includes the balance from current revenue, contributions of public enterprises, additional resource mobilization, plan grants and loans, market borrowings and other miscellaneous capital receipts. The state plans are then approved by the Planning Commission. Thus, in the final analysis, given the amount of central transfers to the states as determined by the Gadgil formula, at the margin it is mainly the own resource position of the states that determines their plan sizes.

Plan assistance does not have any relationship with the investment requirements of the states. Transfers are not directly related to any shortfall in states’ resources, given the required amount of plan investments and own resources reckoned at a standard level of effort. Plan transfers to the states, as well as their grant-loan components, are determined independently of the required plan investments, their sectoral composition, resources available with the states or their fiscal performance. In fact, the grant component of central plan assistance has been kept at 30 per cent because when the Gadgil formula was introduced, the current component of plan outlay was approximately 30 per cent.
Hence, while there were considerable variations in the ratio of current plan expenditures among individual states, the grant-loan mix for plan assistance for the major states has been kept constant. The constancy in the grant portion to all the major states does not take account of the differing repayment abilities of the states. Also, it involves a bias against states with a strategy for development through human capital formation (e.g., education), as against those with an emphasis on material capital formation. In the former, the current expenditure component, according to prevailing budgeting practices, is higher. Since returns on expenditure initially accrue to the individual rather than the government, states with a larger current component of plan expenditures would have as much of an interest liability as states with a larger share of capital expenditures, but with much lower levels of revenue-yielding assets. There may, therefore, be a case for varying the grant component of central plan assistance, depending on the repayment capacity of individual states.

**Assistance for Central Sector and Centrally Sponsored Schemes**

Assistance given to states through central sector and centrally sponsored schemes, constituting about 20 per cent of total transfers, is in some respects the most controversial form of transfers. These transfers are neither based on the recommendations of the Finance Commission, nor determined by the Gadgil formula, but are discretionary. Central government ministries initiate a number of “National Programs,” either by themselves, or at the request of the relevant ministries at the state level. Central sector schemes are assisted entirely by way of central grants and the states merely have the agency function of executing these programs. Centrally sponsored schemes are essentially cost sharing programs, and the share of central assistance is given by way of grants or loans decided for each of the programs. The rationale for introducing these programs is ostensibly to finance activities with a high degree of inter-state spillovers, or are in the nature of merit goods (for example, poverty alleviation and family planning).

Although the major programs on family planning and rural development are well designed, and the transfers are given according to formulas devised by the administering ministries, bureaucratic and political discretion often plays an important role in determining the amount of transfers and the pattern of their distribution. There have also been instances where the Prime Minister has announced the programs in public meetings, leaving the Planning Commission and the relevant ministries to work out details subsequently. If even a few of the programs are determined in an arbitrary and non-transparent manner, well formulated programs under central sector and centrally sponsored schemes also become the subject of doubts about their objectivity and transparency.

These programs have provided the central government with an instrument to actively influence states’ spending. Until 1969, the volume and pattern of assistance to state plan schemes were decided for each project, and the central government did not need such transfers. Once plan assistance was given according to the Gadgil formula, the center introduced these specific purpose transfers and expanded them significantly. These schemes have grown in both volume and number over the years, in spite of states’ objections to their proliferation and the
decision of the National Development Council (NDC) in 1970 to roll into assistance to such schemes 1/6th of central assistance for state plans. At present, there are over 250 centrally sponsored schemes with detailed conditionalities. These conditionalities, such as requirements on staffing patterns, tend to distort the states’ own spending. Also, the proliferation of schemes seemingly has increased the bureaucracy considerably. Therefore, the NDC appointed an investigative committee, which recommended scaling down centrally sponsored schemes. This recommendation, however, has not been acted upon in a serious manner. There is perhaps a strong case for consolidating a number of schemes into specific purpose transfers under broad headings, with greater flexibility given to the states in the use of funds.

Institutional Details

Normative criteria for a successful intergovernmental transfer system are that, besides being equitable and incentive efficient, it should be simple, objective, and transparent. These criteria, in turn, require a proper institutional mechanism. In India, the Constitution attempted to create this through the Finance Commission. The Commission was to be appointed every five years, to take account of the changing needs of the center and the states. The Finance Commission (Miscellaneous) Act also lays down qualifications of the Chairman and Members of the Commission and the presence of a judicial member/chairman in the Commission is supposed to give it an independent, semi-judicial status.

Despite the provision of a specialized, independent and semi-judicial agency, the system of transfers evolved over the years in India has not fulfilled its intended objectives. The design of general and specific purpose transfers falls short of the intention of offsetting the fiscal disabilities of poorer states, and of ensuring minimum standards of services in aided activities. Further, transfers in practice are not simple and transparent; the incentives generated by the system do not promote good fiscal management; and transfers are not well targeted to meet the objective of fiscal equalization. There are a number of institutional reasons for this outcome. First, as already discussed, although the Constitution envisaged rule-based transfers on the recommendation of the Finance Commissions, other developments put a major proportion of transfers outside its purview. Multiple agencies giving transfers in an uncoordinated manner cannot implement singular economic objectives. Furthermore, while the Finance Commission is at least expected to be non-political, the Planning Commission is not. The Gadgil formula used for distributing Plan assistance is determined by consensus in the NDC, where all the states are members. Finally, the centrally sponsored schemes are discretionary, and designed by the central ministries, where many non-economic considerations enter into the distribution mechanism.

Even the Finance Commissions have not functioned well professionally in evolving the transfer system. Lack of permanency in their tenure has impeded the development of a satisfactory methodology for dispensing transfers. Although a small cell has been created in the Finance Ministry, it is ill-equipped to improve the methodology of making projections, estimating fiscal capacities and needs of the states, undertaking analysis of their indebtedness; or to maintain and update the data required for the analysis by the subsequent Commission. Each Commission has to start afresh and, given its time constraints, finds it difficult to conduct the
analysis necessary for making recommendations consistent with overall objectives. Thus, there has been very little improvement in methodology or databases.

The fact that the central government (through the Ministry of Finance) determines the Chairman and Members of the Commission and specifying its terms of reference, raises questions about objectivity and fairness in the minds of state-level decision-makers. This is particularly true when political personalities are appointed to the Commission. Also, the Member-Secretary is always a senior bureaucrat belonging to the Indian Administrative Service, appointed not because of expertise or interest in the subject, but because he or she qualifies to be appointed as a Secretary. Sometimes, mid-way through the Commission’s tenure, the Member-Secretary secures a transfer to a more prestigious posting as a Secretary in an important administrative department, and is replaced by another such bureaucrat. The Commission’s staff also comes chiefly on deputation from various central ministries: many are unfamiliar with the technical details of state finances, intergovernmental transfers, and research methods.

Lack of co-ordination between the Planning and Finance Commissions further adds to the shortcomings of the current institutional arrangement. There have been cases where the Planning Commission set about filling the non-plan gaps of the states in their current accounts, resulting from their non-compliance with the norms set by the Finance Commissions. The states, as noted earlier, submit different projections of revenues and expenditures to the two commissions. The presence of a common member in Planning and Finance Commissions has partly resolved this issue, but the problem of independent treatment of interdependent plan and non-plan sides remains.

VI. Empirical Framework and Results

Our goal is to use a parsimonious framework to try to explain the observed pattern of center-state fiscal transfers. We restrict attention to explicit current transfers, having discussed elsewhere (Rao and Singh, 1998d) implicit transfers through subsidized loans to states. By a parsimonious framework, we mean the use of regressions with just a few key variables that describe the economic, demographic and political characteristics of the states. We begin by describing the data and estimation methodology, then present the results.

Data

We use data on fiscal transfers from the center to the states that has been compiled and cleaned up at the National Institute for Public Finance and Policy, New Delhi. This Institute is primarily responsible for such data compilation and analysis, and we are quite confident about the data quality. There are no missing observations. We use data on 14 major states (these exclude the so-called Special Category states, and the small state of Goa, which was upgraded from Union Territory status relatively recently). The Special Category states are distinguished by being border states, with substantial ethnic and religious differences from the “mainstream” of India. Their exclusion therefore does not imply unimportance. In fact, as the data in the last

---

10 This designation is the highest rung of the civil service.
column of Table 5-3 suggests, they are the clearest illustration of a Rikerian view of federalism. However, there are several wrinkles in considering the special category states (such as differences in cost structures) that suggest a separate analysis with additional data is required. While we have a time series on fiscal transfers that stretches further back, we have, for tractability, restricted our empirical analysis in this paper to the 10-year period from 1983-84 to 1992-93. This NIPFP data set also included figures on State Domestic Products in current and constant prices, and in total and per capita terms. Thus we recovered state population figures from the ratio of per capita and total values for each year and deflators from the ratio of current to constant price figures. These were used to convert the fiscal data to constant price terms (with 1981 as the base year), and to per capita terms wherever required. We describe the fiscal data in more detail below.

We also use data on political characteristics of the states. In particular, we use data on the share of different states' parliamentary representation\(^{11}\) in the ruling party or ruling coalition. This data is constructed from Butler, Lahiri and Roy (1995). The period of estimation included majority Congress governments from 1980 to 1984 and 1985 to 1989, as well as a minority Congress government from 1991 to 1996. From 1989 to 1991 (a period of about a year and a half), there were two Janata Dal minority governments. The existence of 'outside support' for minority governments introduces some noise into using the share variable as a measure of political strength of the state in the central process, but for the present paper we work with this variable.\(^{12}\) We also use data on the control of the various state governments, using this to create a variable which takes the value one if and only if the party at the center and the state level are the same in a particular year, and zero otherwise.\(^{13}\)

We now describe the data on transfers in greater detail. The table below illustrate the tax data we have, using the original (current price) data from Andhra Pradesh for 1983-84. For our present analysis, we consider only the aggregate of shared taxes. Recall that these are centrally collected taxes, which are constitutionally required to be shared with the states. The aggregate share, and the distribution among the states, are determined by the Finance Commission.

---

\(^{11}\) We used figures for the lower house only, since this is the main legislative body. The upper house has limited, though not completely negligible powers.

\(^{12}\) Ideally, we would like to calculate a power index, such as the Banzhaf Index, to measure the political clout of different states in the ruling party/coalition at the center, but such calculations will require implementing a complex computer program, which we have obtained, but not yet tried. Sugata Marjit, in a conversation with the second author, has suggested that he has obtained promising results using representation of states in the central cabinet. This is also very plausible, but we do not have further details on his work. An idea in a similar vein, which is based on the experience of the first author in working with the Finance Commission, is that the state allegiances of the Members of the Commission, in particular the Chairman, have an impact on actual transfers.

\(^{13}\) This data was kindly made available to us by Bhaskar Dutta, who used it in an analysis of the state governments' expenditure patterns (Dutta, 2000). The data is also in Butler, Lahiri and Roy (1995). Again, the existence of coalition governments in states can make the matching variable we use somewhat less reliable.
Total Tax Revenue = 118440.6

Own Tax Revenue + 82352.0

Shared Taxes = 36088.6

Shared Income Taxes + 9069.5

Shared Estate Duty + 107.5

Share of Union Excise Duties = 26911.7

Basic Union Excise Duties + 23294.2

Additional Excise Duties 3617.5

The data on non-tax revenue of the states is also available broken down by categories. The table below illustrates the nature of the original data, also using figures from Andhra Pradesh, again at current prices for 1983-84. The four grant categories are further disaggregated in the original data, but we do not present that disaggregated data here.

Total Non Tax Revenue = 57966.50

Total Own Non Tax Revenue + 30942.44

Grants from Central Government = 27024.06

Non Plan Grants + 6317.80

Grants for State Plan Schemes + 7862.07

Grants For Central Plan Schemes + 3108.55

Grants For Centrally Sponsored Schemes 9735.65
For the empirical analysis in this paper, we aggregate transfers into three categories, as indicated below:

Statutory transfers = Shared Taxes + Non Plan Grants
Grants for State Plan Schemes
Discretionary transfers = Grants For Central Plan Schemes + Grants For Centrally Sponsored Schemes

The sum of these three categories constitutes Total Transfers. We also run regressions using this total variable. Disaggregation of categories of transfers beyond that assumed in this paper is certainly possible, and we shall explore that in future work. Summary statistics for the data are provided in Table 5-7. The means in the table are unweighted means of the 140 observations. Correlations for the independent variables are also calculated similarly. The summary statistics suggest that, in general, neither lack of variation nor high correlation between independent variables is likely to be a problem.

**Estimation Methodology**

We used the LIMDEP7 program to estimate fixed effects models for various specifications. We report selected results in detail in this paper, and briefly discuss other specifications. All regressions were run alternatively using the three transfer components, and their total, as the dependent variables, in constant price, per capita terms. The independent variables used were State Domestic Product at 1981 prices (SDP81), per capita constant price SDP (SDPPC81), population (POPN), the proportion of the ruling party's Members of Parliament (lower house only) coming from a particular state (PROP), and the variable measuring whether the same party was in power at the center and the state level (MATCH). Lags of the latter two variables were also tried, with, for example, a three-year lagged variable being denoted PROPNLAG3 and MATCHLAG3 respectively. We found the results with the lagged variables to be more plausible, as we discuss below, and only those are reported. In addition to these independent variables, state fixed effects were also included. The model without state fixed effects was always rejected in the standard F-tests automatically carried out by LIMDEP, and therefore those results are not reported. We tried three specifications: linear, loglinear, and translog. In each case the political variables were unchanged. All estimations were carried out using the White heteroscedasticity-corrected variance covariance matrix.

One lacuna in the regressions reported here is that we were unable to test the fixed effects specification against a random effects specification, using Hausman's (1978) standard test. LIMDEP was never able to compute the required statistic, although it did produce the random effects estimates. In general, the random effects estimates were somewhat different from the fixed effects estimates, and seemed less stable and less plausible. In general, we believe that the random effects model requires independence assumptions that are unlikely to be satisfied (as was the case in Hausman's original illustration of the test), and we therefore present only the fixed effects results in this paper.14

---

14 The random effects results are available from the authors on request. They do display somewhat similar impacts on transfers of the political and other variables.
Linear Specification Results

Table 5-8 presents results for the linear specification, for each of the four dependent variables (the three components of transfers, as well as their sum). The first regression, for purposes of illustration, reports the regression without the two explicit ‘political’ variables. The dependent variable here is statutory transfers. In general, we found that the coefficients of SDP, SDPPC and POPN were not very sensitive to inclusion of the political variables. Therefore, except for the illustrative regression (1), all other regressions reported include the two political variables. As noted, we found that a lag of three years for the political variables gave reasonable results, and we imposed this lag for all the regressions reported. The reasoning behind using lags is that actual transfers would be heavily influenced by factors determined in advance by the Finance Commission, and political impacts would therefore show up with a lag. A rationale for the particular lag can be sought in the five year cycle of Finance Commission and Planning Commission awards, with three years representing an approximate ‘average’ lag, though this is only a rough intuition. The results with the lagged political variables are reported for the linear regressions numbered (2)-(5) of Table 5-8.

Note that, since the three independent variables SDP, SDPPC and POPN are multiplicatively related\(^5\), the coefficients of these variables can not directly give us marginal impacts of changes in state characteristics. The marginal effects at the unweighted means of the data are therefore also reported, in Table 5-9. We use the point estimates to calculate the marginal effects, even though some estimated coefficients are statistically ‘insignificant’ at conventional levels. In every case, except for regression (3), the ‘t’-ratios are at least one, and so the point estimates are not only statistically ‘best’, but are also reasonably precise. For regression (3), with grants for state plan schemes as the dependent variable, none of the three coefficients are even close to statistically significantly different from zero at conventional levels, and we have omitted the calculation of marginal effects for this regression from Table 5-9. For the other regressions, since the marginal effects are linear functions of the coefficients, we are able to calculate the usual t-statistics, and these are reported in Table 5-9 as well.

We discuss the coefficients and marginal effects for each of the regressions (2)-(6) in turn. We have also presented the fixed effects for regressions (2)-(6), in Table 5-10, and we examine them for all four regressions together, once we have presented the discussion of the coefficients and marginal effects.

The regression for statutory transfers per capita, (2), has only the coefficients for SDP per capita and population statistically significantly different from zero. However, the other coefficients are not estimated too imprecisely, and the overall fit is reasonable, though much of the explanatory power comes from the fixed effects. The main story is to be found in the estimated marginal effects in Table 5-9. There we see that the effect of changes in SDP per capita on the per capita statutory transfer, controlling either for economic size (SDP) or demographic size (population), is surprisingly not statistically significant, though in the first case

\(^{15}\) Because of the units we have used for the three variables, the relationship is SDP81 = POPN*SDPPC81*10.
it has the "right" negative sign. The latter is consistent with equalization objectives and the evidence from simple correlations presented in the last section. On the other hand, higher SDP is associated with higher per capita statutory transfers. Thus, whether we control for the per capita product of the state, or for its population, a state which is economically more important, as measured by the size of economic activity in the state, receives higher per capita transfers. This result is not an obvious outcome of the complex institutional process of making transfers, and it is our first important observation from our regressions. Finally, population also has a positive effect on per capita transfers: a more populous state receives more per capita, compared either to a state with the same per capita product or to a state with the same total domestic product (i.e., with a higher per capita product).

The political variables in regression (2) are statistically insignificant. We tried several other combinations and lags of the two political variables in the statutory transfers regression. In no case were the coefficients of the two political variables statistically significantly different from zero. As noted, we report only the results for the specification where the variables were lagged by three years. However, we may interpret the population variable as capturing some political effects: in a democracy, the demographic size of a state may be an indicator of its political influence.

Regression (3) in Table 5-8 presents results for the case where the dependent variable is per capita grants for state plan schemes. Only the matching variable is statistically significant, with a lag of three years. The effect is economically important as well. The mean per capita grant for the sample is Rs. 13.97, and the average effect of the state being ruled by the same party as the center is that the per capita grant in this category is higher by Rs. 4.78. The interpretation of this regression is quite plausible: states propose these schemes to the Planning Commission, and states that are ruled by the party in power will tend to receive higher per capita grants. The other political variable is insignificant, as are the measures of state economic activity and demographics (making calculation of those marginal effects moot). Alternative specifications did not change these latter results. The only noteworthy feature was that when the matching variable was included without a lag, it was negative and statistically significant at the 10% level. We do not have a plausible explanation for this result: such a switch in sign was not observed when lags were varied in the statutory transfers regression. The results for the grants for state plan schemes regression, when contrasted with those for statutory transfers, show very clearly that different components of transfers are determined by very different factors. This regression's explanatory power is also considerably lower, suggesting that unobserved factors are at work for this component of transfers.

Regression (4) in Table 5-8 presents the results where per capita discretionary (as defined by us) grants are the dependent variable. Neither of the political variables is significant, although when we included only the current MATCH variable, it was again, surprisingly, negative, and significant at the 10% level. The marginal effects, reported in Table 5-9, are quite similar in size and magnitude to those for statutory transfers, even though the institutional determination of the two different categories is very different. While the coefficients are mostly somewhat smaller in magnitude, it must be noted that transfers in the discretionary category are less than one-third of
those in the statutory category. Thus the marginal impacts are proportionately higher. For example, an increase in SDP by 100,000 units (Rs. 1 trillion) is estimated to increase statutory per capita transfers by about Rs. 4.40, which is 7.6% of the mean transfer in this category. The estimated effect on discretionary transfers is about Rs. 2.80, which is 15% of the mean transfer in this category. Similarly, population has greater proportionate effects for per capita discretionary transfers as compared to statutory transfers. Perhaps discretionary transfers are more politically useful in larger states, irrespective of whether the ruling party currently controls the state or not. Also, the statistical significance levels for the marginal effects are somewhat greater in the case of discretionary transfers, than for statutory transfers. One surprise in the results is the statistically significant negative effect on discretionary transfers per capita of higher per capita SDP, keeping total SDP constant. Our hypothesis would have been that discretionary transfers were less likely to display such equalization effects.

Regression (5) in Table 5-8 presents results for total transfers per capita as the dependent variable. The coefficients of the demographic and income variables, and the resulting marginal effects shown in Table 5-9, are not dissimilar to those for statutory and discretionary transfers. The lagged matching variable appears to reflect the impact on grants for state plan schemes that was discerned in the regression for that component of transfers, and is very close to being significant at the 10% level. The lagged variable measuring the proportionate importance of the different states in the ruling party's parliamentary strength is the closest it comes to statistical significance in any of the linear regressions. In sum, the total transfers regression appears to reflect quite clearly the determinants of the three components of transfers. Of these three components, statutory and discretionary transfers are surprisingly similar in their determinants, while grants for state plan schemes are clearly influenced by a different set of factors, which are not being well captured in this empirical exercise.

Table 5-10 also shows the fixed effects coefficients for the 14 states, for each of the regressions, (2)-(5). The coefficients for the statutory transfers regression display considerable variation, ranging from -96 to 33. Several of the fixed effect coefficients (particularly those that are larger in magnitude) are statistically significantly different from zero. The variation in the fixed effects suggests, of course, that factors missing from the regressions are important determinants of transfers. One possibility is that the poverty rate is an important missing variable, since Bihar, Madhya Pradesh and Uttar Pradesh all have negative fixed effects that are large in absolute value. Furthermore, Maharashtra, also with a sizable negative fixed effect, while a high-income state, has a poverty ratio that is relatively high for its SDP per capita, the average being skewed by the large financial center, Mumbai. However, higher poverty ratios should be reflected in fixed effects that are positive, or less negative, rather than the pattern that is observed. Furthermore, Rajasthan, the fourth ‘BIMARU’ state16 does not fit the pattern suggested by the relative poverty explanation (or being part of the Hindi heartland, which also fits the BIMARU states). Neither does the fixed effect for Orissa, another poor state, fit this explanation.

---

16 The word ‘bimar’ means ‘ill’ in Hindi. The term ‘BIMARU’ comes from the beginning letters of Bihar, MP, Rajasthan and UP.
Another possibility is that nonlinearities with respect to the impact of the independent variables can account for the variation in the fixed effects. In particular, the fixed effect coefficients seem to be larger in magnitude for states with larger populations (e.g., Bihar, Maharashtra and Uttar Pradesh). Therefore we did try to allow for nonlinearities by including, alternatively the square and the square root of population, in addition to the other independent variables. However, this resulted in severe multicollinearity, with unstable and insignificant coefficient estimates, as well as other estimation problems in some cases. Furthermore, the fixed effect coefficients continued to exhibit considerable variation across states. Explaining the fixed effects therefore remains an issue, which we return to later in the paper.

The fixed effects for the grants for state plan schemes regressions (column (3) of Table 5-10) are smaller than for the statutory transfers regressions (and now none are statistically significantly different from zero), but almost proportionately so. They display some of the same patterns across states as in the statutory transfers regression. The fixed effects for discretionary transfers also show considerable variation across states. Now their magnitudes are quite large, compared to the fixed effects for the larger category of statutory transfers, and they are mostly statistically significant. The pattern of fixed effects for discretionary transfers does show some differences, compared to the other two categories of transfers. For example, the fixed effects for Tamil Nadu and West Bengal are relatively low (more negative) than for statutory transfers or grants for state plan schemes. In fact, the differences in the fixed effects mark the greatest difference between the statutory and discretionary transfers regressions, rather than the economic, demographic and measured political factors.

Loglinear Specification Results

In order to explore the sensitivity of our results to different specifications, we next estimated a loglinear specification. The results for these are reported in Table 5-11, with fixed effects given in Table 5-12. Since the logarithm of SDP is the sum of the logarithms of population and per capita SDP, it is omitted from these regressions. The coefficients of SDP per capita and population are now the elasticities of per capita transfers with respect to these variables, keeping the other variable constant. Since $\ln SDPPC = \ln SDP - \ln POPN^{17}$, we can substitute this into the equation to derive the other elasticities. The elasticity with respect to SDP keeping population constant is the same as the elasticity with respect to per capita SDP (keeping population constant) in this model, while the elasticity with respect to per capita SDP keeping SDP constant is the difference in the coefficients (since $\ln POPN = \ln SDP - \ln SDPPC$). The two explicit political variables were included in the regression without transformation.

The significance of the variable that measures states' influence in the national parliament (PROPNI) changes when we try loglinear regressions. In the logarithmic regressions, the variable PROPNLAG3 now has a positive and significant effect for statutory and for total transfers. This result is again somewhat surprising, since one would have hypothesized that discretionary transfers were more subject to these kinds of influence. The coefficient of this variable is also

---

17 There is also a constant term, reflecting the difference in units explained in footnote 15, but this does not affect the elasticity relationships.
positive, though statistically insignificant, for the other two components of total transfers. The variable that measures whether there is a match between the national and state government parties is statistically significant only for grants for state plan schemes, as was the case for the linear specifications.

As noted, the coefficients of the population and per capita SDP variables are now elasticities, and the elasticity with respect to the latter, keeping population constant, is positive (i.e., counter to equalizing objectives) in three of the cases, surprisingly excepting discretionary transfers. However, the elasticity with respect to per capita SDP is not statistically significant in the case of grants for state plan schemes. Also, in all four cases, this elasticity is negative when the state fixed effects are omitted (results not reported here). This again points to the role of state fixed effects as capturing some unaccounted-for economic criterion. The state fixed effects in Table 5-12 display patterns similar to those for the linear regressions in Table 5-10.

One noteworthy result is the large positive coefficient of population in the case of discretionary transfers (keeping per capita SDP constant). This is consistent with the point made for the linear regressions, that population may be an indicator of political influence, solely due to the size of the state, and irrespective of its economic base or contribution to the ruling group in parliament. Finally, the elasticity of per capita transfers with respect to economic size, as measured by SDP, is generally positive in the four logarithmic regressions, since it is either the coefficient of SDP per capita (if population is held constant), or the coefficient of population (if per capita SDP is held constant. The impact is largest for the discretionary transfers regression. These results are therefore consistent with the linear regressions.

In broad terms, therefore, the loglinear regressions are consistent with the linear specification results. There is some evidence in the loglinear regressions (perhaps slightly stronger than in the linear case) that political variables matter. This effect arises in the expected way in the regression explaining grants for state plan schemes, but in an unexpected manner in the case of statutory transfers, which one would have hypothesized were less susceptible to political influence than what we have characterized as discretionary transfers. Even in the loglinear case, there are definite patterns in the fixed effects that suggest that unexplained factors might be captured in those variations. It is possible, however, that the variation in fixed effects is a result of misspecification of the regressions. In particular, neither the linear nor the loglinear specification may capture the effect of the complicated formulas that are used for statutory transfers. We therefore also estimated a more flexible functional form, the translog specification described next.

**Translog Specification Results**

The translog specification is a well-known approximation to a general functional form, based on the idea of a Taylor series expansion of a function. The translog specification results are presented in Table 5-13. Clearly some of the multicollinearity that made the linear-quadratic estimates impossible to calculate is present in the translog estimates. For example, population is now no longer statistically significant, when it enters the regression through three different
variables. However, the cases of statistical significance of the political variables that were present in the loglinear case survive this generalized specification. There is some evidence of nonlinearities in the relationship among the logged variables, indicating that the elasticities of response of transfers to demographic and economic variables are not constant.

It is possible to re-estimate the translog specification in a restricted form, with restrictions varying for each equation, and some preliminary estimates (not reported here) were reasonable, restoring, for example, the statistically significant coefficients on population. However, the main point we wish to emphasize here is that the fixed effects, reported in Table 5-14, are now almost uniform for every one of the regressions in the translog specification. Thus it is at least possible that nonlinearities in the response of per capita transfers to demographic and economic characteristics of the states explain some of the pattern of the fixed effects in the earlier regressions, in addition to the omission of idiosyncratic political and economic factors.

Summary

Overall, the regressions suggest that, even with a very simple specification, there is some evidence for the importance of variables that may proxy bargaining power of the components of the Indian federal system. This conclusion is based on the positive estimated effect of economic and demographic size of the states on both statutory and discretionary transfers per capita, and of the lagged effect of a match between the state and central ruling parties on grants for state plan schemes. In the loglinear case, there is also a positive effect of the proportion of ruling party/coalition MPs on per capita statutory transfers, again with a lag. A cautionary note, besides the general problem of potential fragility of econometric results such as these, is in the limited explanatory power of the independent variables. Most of the explained variance in the regressions reported here is due to the state fixed effects. We have suggested that there may be patterns in these state fixed effects that can be captured by other measurable political and economic characteristics of the states. Examples of such variables include the presence of prominent parliamentary party members with important ministries in their charge, poverty ratios, degree of urbanization, and level of infrastructure. Including such variables may improve the explanatory power of our regressions, and will provide a robustness check on the initial results presented here.

VII. Conclusion

Much of our paper has involved a critical review of how the features of Indian federalism determine the levels and composition of fiscal transfers from the central government to the state governments. The length of our review was necessitated partly by the complexity of the arrangements that have evolved for such transfers over the decades since India’s independence. However, the chief contribution of our paper is in the admittedly preliminary empirical exercise reported in the previous section. There, rather than engage in yet another normative dissection of the details of India’s system of center-state transfers, we have attempted to cut through the
institutional thickets with a simple positive empirical model of the main components of vertical transfers.

Our motivation is in the view of a federal system as a constitutional or political bargain. Even though India was not formed out of an explicit bargaining process (except to some extent with respect to the inclusion of the princely states at the time of independence), the perspective of bargaining is commonly applied informally to resource sharing among the different constituent governments. The states, while they have not had sovereign status, and, constitutionally speaking, exist at the pleasure of the central government, represent real and significant political groupings, based on language and culture. We would argue that they are the subnational political units that matter above all, more so than caste or class. Therefore, center-state transfers in India, which are large in relative and absolute terms, provide a natural data set with which to test hypotheses on functioning of a federal system as an ongoing political bargain. As we outlined in Section II, we simultaneously tackle the distributive issues that arise due to the possibility of secession, as well as those that are part of the normal politics of ongoing governance.

After providing a detailed discussion of India’s institutions and experience with respect to vertical transfers, we therefore proceeded to run some explanatory regressions. As far as we are aware, this is the first such exercise to be attempted for India. Given the heterogeneity of methods of transfer, we grouped transfers in this initial analysis into two broad categories, which we termed statutory and discretionary transfers, leaving separate a third category, grants for state plan schemes. Surprisingly, results for statutory and discretionary transfers were broadly similar in the linear regressions, despite the very different institutional mechanisms governing them. On the other hand the factors governing grants for state plan schemes seemed to be quite different, and tied to political considerations in a plausible way. The loglinear specifications provide further evidence for the hypothesis that the political and economic importance of the states has a positive influence on per capita transfers.

We have noted, in Section II, as well as at the end of the last section, the need to incorporate further variables into our empirical analysis. To this we can add plans to investigate alternative specifications, and longer time spans. We may also note again that explicit transfers are only one component of a complex process of distribution of the surplus: tax rates, the location of economic activity, the states’ own fiscal behavior, and other variables also enter. To the extent that we can measure these factors, they represent an interesting possibility for future work. Another important future task is to compare our results for India with the fast-growing analytical and empirical work on federal transfers in other countries.

---

18 These results and ideas were first presented in preliminary form at the NIPFP in September 1998. As noted earlier, we recently became aware that Sugata Majit has undertaken a similar empirical exercise, though we have not yet seen his work.

Our emphasis in the analysis is positive, to examine the overall outcomes emerging from a complex and heterogeneous set of institutions and motivations. However, if our analysis can show that the outcomes exhibit patterns, predictable or unexpected, our work may ultimately aid in designing a more effective set of institutions for intergovernmental transfers in India. Alternatively or additionally, the ultimate conclusion may also be that a system with large vertical transfers is inevitably subject to political pressures and unintended effects, implying a need to reconfigure the underlying tax assignments to achieve a better match with expenditure responsibilities at different levels of government.

References


Table 5-1: Revenue Receipts of the Central and State Governments

(Per cent)

<table>
<thead>
<tr>
<th>Items of Revenue</th>
<th>Revenue Share 1985-86</th>
<th>Revenue Share 1990-91</th>
<th>Revenue Share 1997-98</th>
<th>Per cent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Center</td>
<td>States</td>
<td>Center</td>
<td>States</td>
</tr>
<tr>
<td>A. Tax Revenue (a+b+c)</td>
<td>49.0</td>
<td>51.0</td>
<td>49.1</td>
<td>50.9</td>
</tr>
<tr>
<td>a. Exclusive Central Taxes</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td>1. Corporation Tax</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td>2. Custom Duties</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td>3. Other</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td>b. Exclusive State Taxes</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>1. State Excise Duties</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>2. Sales Taxes</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>3. Taxes on Transport</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>4. Other</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>c. Shared Taxes</td>
<td>51.6</td>
<td>48.4</td>
<td>51.4</td>
<td>48.6</td>
</tr>
<tr>
<td>1. Personal Income Tax</td>
<td>26.5</td>
<td>73.5</td>
<td>23.4</td>
<td>76.6</td>
</tr>
<tr>
<td>2. Union Excise Duty</td>
<td>56.6</td>
<td>43.4</td>
<td>57.5</td>
<td>42.5</td>
</tr>
<tr>
<td>B. Non-tax Revenue</td>
<td>62.1</td>
<td>37.9</td>
<td>54.3</td>
<td>45.7</td>
</tr>
<tr>
<td>1. Net Contribution from Public Enterprises</td>
<td>-875.9</td>
<td>975.9</td>
<td>-288.1</td>
<td>388.1</td>
</tr>
<tr>
<td>2. Administrative Receipts</td>
<td>20.8</td>
<td>79.2</td>
<td>33.8</td>
<td>66.2</td>
</tr>
<tr>
<td>2. Interest Receipts</td>
<td>66.6</td>
<td>33.4</td>
<td>59.6</td>
<td>40.4</td>
</tr>
<tr>
<td>3. External Grants</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td>C. Grants to States</td>
<td>100.0</td>
<td>-</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td>D. Total Revenue Accrual</td>
<td>38.2</td>
<td>61.8</td>
<td>37.7</td>
<td>62.3</td>
</tr>
<tr>
<td>E. Total Revenue Collections</td>
<td>65.6</td>
<td>34.4</td>
<td>63.9</td>
<td>36.1</td>
</tr>
</tbody>
</table>

Table 5-2: Trends in Vertical Fiscal Imbalance

<table>
<thead>
<tr>
<th>Period</th>
<th>Per cent of States' own current revenue to total current revenue</th>
<th>Per cent of States current expenditure to total current expenditure</th>
<th>Per cent of States' own current revenue to States' current expenditure*</th>
<th>Per cent of States' expenditure* to total expenditure*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1955-56</td>
<td>41.2</td>
<td>59.0</td>
<td>68.9</td>
<td>61.7</td>
</tr>
<tr>
<td>1960-61</td>
<td>36.6</td>
<td>59.9</td>
<td>63.9</td>
<td>56.8</td>
</tr>
<tr>
<td>1965-66</td>
<td>32.6</td>
<td>55.6</td>
<td>63.5</td>
<td>53.3</td>
</tr>
<tr>
<td>1970-71</td>
<td>35.5</td>
<td>60.2</td>
<td>60.6</td>
<td>53.9</td>
</tr>
<tr>
<td>1975-76</td>
<td>33.5</td>
<td>55.1</td>
<td>70.4</td>
<td>47.6</td>
</tr>
<tr>
<td>1980-81</td>
<td>35.6</td>
<td>59.6</td>
<td>60.1</td>
<td>56.0</td>
</tr>
<tr>
<td>1985-86</td>
<td>35.5</td>
<td>56.0</td>
<td>57.7</td>
<td>52.6</td>
</tr>
<tr>
<td>1990-91</td>
<td>36.6</td>
<td>55.2</td>
<td>53.5</td>
<td>53.1</td>
</tr>
<tr>
<td>1995-96</td>
<td>39.6</td>
<td>57.2</td>
<td>58.3</td>
<td>56.7</td>
</tr>
<tr>
<td>1996-97</td>
<td>37.3</td>
<td>58.3</td>
<td>54.2</td>
<td>88.5</td>
</tr>
<tr>
<td>1997-98</td>
<td>39.5</td>
<td>57.6</td>
<td>55.5</td>
<td>69.4</td>
</tr>
</tbody>
</table>

* Current + capital expenditures.

Source: Public Finance Statistics, Ministry of Finance, Government of India (relevant years).
Table 5-3: Revenues and Expenditures of the States - 1993-94

<table>
<thead>
<tr>
<th>I. Major States</th>
<th>Per capita SDP (Rupees)*</th>
<th>Poverty ratio (per cent)**</th>
<th>Per capita own revenue (Rupees)</th>
<th>Own revenue as percentage of SDP</th>
<th>Per capita current spending (Rupees)</th>
<th>Per cent of revenue to current spending</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. High Income States</td>
<td>10211</td>
<td>32.7</td>
<td>1278.63</td>
<td>12.5</td>
<td>1680.87</td>
<td>76.1</td>
</tr>
<tr>
<td>1. Gujarat</td>
<td>7600</td>
<td>32.3</td>
<td>1233.82</td>
<td>16.2</td>
<td>1601.92</td>
<td>77.0</td>
</tr>
<tr>
<td>2. Goa</td>
<td>11658</td>
<td>23.4</td>
<td>2632.4</td>
<td>22.6</td>
<td>3499.84</td>
<td>75.2</td>
</tr>
<tr>
<td>3. Haryana</td>
<td>10359</td>
<td>16.6</td>
<td>1680.61</td>
<td>16.2</td>
<td>1951.10</td>
<td>86.1</td>
</tr>
<tr>
<td>4. Maharashtra</td>
<td>10984</td>
<td>40.1</td>
<td>1213.82</td>
<td>11.1</td>
<td>1578.65</td>
<td>76.8</td>
</tr>
<tr>
<td>5. Punjab</td>
<td>12319</td>
<td>12.7</td>
<td>1214.61</td>
<td>9.9</td>
<td>1915.74</td>
<td>63.4</td>
</tr>
<tr>
<td>B. Middle Income States</td>
<td>6661</td>
<td>38.2</td>
<td>765.59</td>
<td>11.5</td>
<td>1238.84</td>
<td>61.8</td>
</tr>
<tr>
<td>1. Andhra Pradesh</td>
<td>6651</td>
<td>27.2</td>
<td>744.43</td>
<td>11.2</td>
<td>1151.37</td>
<td>64.7</td>
</tr>
<tr>
<td>2. Karnataka</td>
<td>7029</td>
<td>38.1</td>
<td>970.42</td>
<td>13.9</td>
<td>1325.3</td>
<td>73.2</td>
</tr>
<tr>
<td>3. Kerala</td>
<td>6242</td>
<td>32.1</td>
<td>884.13</td>
<td>14.2</td>
<td>1422.7</td>
<td>62.1</td>
</tr>
<tr>
<td>4. Tamil Nadu</td>
<td>7352</td>
<td>45.1</td>
<td>960.32</td>
<td>13.1</td>
<td>1527.72</td>
<td>62.9</td>
</tr>
<tr>
<td>5. West Bengal</td>
<td>6055</td>
<td>44.0</td>
<td>447.78</td>
<td>7.4</td>
<td>959.89</td>
<td>46.7</td>
</tr>
<tr>
<td>C. Low Income States</td>
<td>4674</td>
<td>46.1</td>
<td>438.97</td>
<td>9.4</td>
<td>969.79</td>
<td>45.3</td>
</tr>
<tr>
<td>1. Bihar</td>
<td>3650</td>
<td>53.4</td>
<td>288.13</td>
<td>7.9</td>
<td>800.22</td>
<td>36.0</td>
</tr>
<tr>
<td>2. Madhya Pradesh</td>
<td>5485</td>
<td>43.4</td>
<td>585.01</td>
<td>10.7</td>
<td>1077.75</td>
<td>54.3</td>
</tr>
<tr>
<td>3. Orissa</td>
<td>4726</td>
<td>55.6</td>
<td>384.07</td>
<td>8.1</td>
<td>1048.61</td>
<td>36.6</td>
</tr>
<tr>
<td>4. Rajasthan</td>
<td>5220</td>
<td>34.6</td>
<td>673.13</td>
<td>12.9</td>
<td>1267.67</td>
<td>53.1</td>
</tr>
<tr>
<td>5. Uttar Pradesh</td>
<td>4744</td>
<td>42.0</td>
<td>401.48</td>
<td>8.5</td>
<td>911.47</td>
<td>44.1</td>
</tr>
<tr>
<td>II. Special Category States</td>
<td>5607</td>
<td>29.7</td>
<td>437.56</td>
<td>7.8</td>
<td>1939.48</td>
<td>22.6</td>
</tr>
<tr>
<td>1. Arunachal Pradesh</td>
<td>7904</td>
<td>37.5</td>
<td>964.11</td>
<td>12.2</td>
<td>4330.91</td>
<td>22.3</td>
</tr>
<tr>
<td>2. Assam</td>
<td>5916</td>
<td>36.8</td>
<td>530.94</td>
<td>6.9</td>
<td>1223.0</td>
<td>33.2</td>
</tr>
<tr>
<td>3. Himachal Pradesh</td>
<td>6519</td>
<td>15.5</td>
<td>693.16</td>
<td>10.6</td>
<td>2489.53</td>
<td>27.8</td>
</tr>
<tr>
<td>5. Manipur</td>
<td>5362</td>
<td>32.9</td>
<td>238.37</td>
<td>4.5</td>
<td>2243.74</td>
<td>10.6</td>
</tr>
<tr>
<td>6. Meghalaya</td>
<td>5519</td>
<td>34.6</td>
<td>399.53</td>
<td>7.2</td>
<td>2528.58</td>
<td>15.8</td>
</tr>
<tr>
<td>7. Mizoram</td>
<td>6599</td>
<td>32.5</td>
<td>462.84</td>
<td>6.7</td>
<td>5399.42</td>
<td>8.6</td>
</tr>
<tr>
<td>8. Nagaland</td>
<td>5870**</td>
<td>34.9</td>
<td>311.42</td>
<td>4.8</td>
<td>5015.25</td>
<td>6.2</td>
</tr>
<tr>
<td>9. Sikkim</td>
<td>5416**</td>
<td>34.7</td>
<td>886.70</td>
<td>15.6</td>
<td>3916.40</td>
<td>22.2</td>
</tr>
<tr>
<td>10. Tripura</td>
<td>3781</td>
<td>36.8</td>
<td>202.16</td>
<td>5.1</td>
<td>2089.88</td>
<td>9.7</td>
</tr>
<tr>
<td>All States</td>
<td>6287</td>
<td>39.3</td>
<td>653.5</td>
<td>10.4</td>
<td>1158.24</td>
<td>56.4</td>
</tr>
</tbody>
</table>

Note: SDP = State Domestic Product.
*Quick Estimates of Govt. of India. **Estimate made by the Expert Committee (India, 1993).
Table 5-4: Central Transfers to States

<table>
<thead>
<tr>
<th>Years</th>
<th>Per Capita Transfers 1981-82 Rupees</th>
<th>Transfers as Percentage of GDP</th>
<th>Transfers as Percentage of Central Revenues</th>
<th>Transfers as Percentage of State Current Revenues</th>
<th>Transfers as Percentage of Total State Expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975-76</td>
<td>77.36</td>
<td>3.67</td>
<td>31.8</td>
<td>38.64</td>
<td>44.80</td>
</tr>
<tr>
<td>1980-81</td>
<td>105.37</td>
<td>4.84</td>
<td>34.8</td>
<td>43.81</td>
<td>47.50</td>
</tr>
<tr>
<td>1985-86</td>
<td>151.54</td>
<td>5.55</td>
<td>40.98</td>
<td>45.62</td>
<td>46.42</td>
</tr>
<tr>
<td>1990-91</td>
<td>179.98</td>
<td>5.20</td>
<td>40.02</td>
<td>44.34</td>
<td>34.22</td>
</tr>
<tr>
<td>1991-92</td>
<td>184.03</td>
<td>5.35</td>
<td>39.65</td>
<td>42.33</td>
<td>34.06</td>
</tr>
<tr>
<td>1992-93</td>
<td>191.04</td>
<td>5.45</td>
<td>40.64</td>
<td>44.16</td>
<td>35.61</td>
</tr>
<tr>
<td>1993-94</td>
<td>194.12</td>
<td>5.33</td>
<td>44.07</td>
<td>42.36</td>
<td>35.38</td>
</tr>
<tr>
<td>1994-95</td>
<td>180.31</td>
<td>4.76</td>
<td>39.09</td>
<td>37.73</td>
<td>30.37</td>
</tr>
<tr>
<td>1995-96</td>
<td>185.39</td>
<td>4.66</td>
<td>39.06</td>
<td>38.37</td>
<td>30.85</td>
</tr>
</tbody>
</table>

### Table 5-5: Current Transfers from the Center to the States

(Rs. Billion, in current rupees)

<table>
<thead>
<tr>
<th>Plan Periods/Years</th>
<th>Finance Commission Transfers</th>
<th>Plan Grants</th>
<th>Other Grants</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tax Devolution</td>
<td>Grants</td>
<td>Total</td>
<td>State Plan Schemes</td>
</tr>
<tr>
<td>Fourth Plan (1969-74)</td>
<td>45.6 (54.2)</td>
<td>8.6 (10.2)</td>
<td>54.2 (64.6)</td>
<td>10.8 (12.8)</td>
</tr>
<tr>
<td>Fifth Plan (1974-79)</td>
<td>82.7 (50.2)</td>
<td>28.2 (17.1)</td>
<td>110.9 (67.3)</td>
<td>29.1 (17.7)</td>
</tr>
<tr>
<td>Sixth Plan (1980-85)</td>
<td>237.3 (57.0)</td>
<td>21.4 (5.1)</td>
<td>258.7 (62.1)</td>
<td>73.8 (17.7)</td>
</tr>
<tr>
<td>Seventh Plan (1985-90)</td>
<td>494.6 (54.2)</td>
<td>62.7 (6.9)</td>
<td>557.4 (61.0)</td>
<td>155.2 (17.1)</td>
</tr>
<tr>
<td>Eighth Plan:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-92</td>
<td>172.0 (52.2)</td>
<td>34.5 (10.5)</td>
<td>206.4 (62.7)</td>
<td>57.2 (14.2)</td>
</tr>
<tr>
<td>1992-93</td>
<td>205.2 (53.5)</td>
<td>26.4 (6.9)</td>
<td>231.7 (60.4)</td>
<td>78.4 (20.4)</td>
</tr>
<tr>
<td>1993-94</td>
<td>223.9 (51.4)</td>
<td>20.7 (4.8)</td>
<td>244.6 (56.1)</td>
<td>107.7 (24.7)</td>
</tr>
<tr>
<td>1994-95</td>
<td>248.5 (52.6)</td>
<td>24.3 (5.2)</td>
<td>272.8 (57.8)</td>
<td>99.0 (21.0)</td>
</tr>
<tr>
<td>1995-96</td>
<td>290.5 (58.0)</td>
<td>39.7 (7.9)</td>
<td>330.2 (66.0)</td>
<td>81.4 (16.3)</td>
</tr>
<tr>
<td>Ninth Plan:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996-97</td>
<td>350.4 (60.2)</td>
<td>36.0 (6.2)</td>
<td>386.4 (66.4)</td>
<td>116.9 (20.1)</td>
</tr>
<tr>
<td>1997-98</td>
<td>400.5 (59.1)</td>
<td>25.5 (3.8)</td>
<td>426.0 (62.9)</td>
<td>128.8 (19.0)</td>
</tr>
<tr>
<td>1998-99</td>
<td>408.9 (57.6)</td>
<td>20.7 (2.9)</td>
<td>429.6 (60.5)</td>
<td>153.5 (21.6)</td>
</tr>
</tbody>
</table>

Note: Figures in parenthesis are percentages total transfers.

Table 5-6: Equalizing Effect of Transfers

<table>
<thead>
<tr>
<th>Transfers</th>
<th>Correlation coefficients with per capita SDP</th>
<th></th>
<th></th>
<th>Income elasticities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared Taxes</td>
<td>-0.167</td>
<td>-0.706**</td>
<td>-0.849**</td>
<td>-0.809**</td>
</tr>
<tr>
<td>Non-plan grants</td>
<td>-0.240</td>
<td>-0.289</td>
<td>-0.110</td>
<td>-0.286</td>
</tr>
<tr>
<td>Total Finance Commission</td>
<td>-0.272</td>
<td>-0.551*</td>
<td>-0.664**</td>
<td>-0.765*</td>
</tr>
<tr>
<td>transfers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plan grants-State plan schemes</td>
<td>-0.263</td>
<td>-0.524*</td>
<td>-0.010</td>
<td>-0.425**</td>
</tr>
<tr>
<td>Plan grants-Central schemes</td>
<td>0.342</td>
<td>-0.101</td>
<td>-0.162</td>
<td>-0.278</td>
</tr>
<tr>
<td>Total plan grants</td>
<td>0.091</td>
<td>-0.327</td>
<td>-0.092</td>
<td>-0.417</td>
</tr>
<tr>
<td>Gross current transfers</td>
<td>-0.194</td>
<td>-0.519*</td>
<td>-0.663**</td>
<td>-0.716**</td>
</tr>
</tbody>
</table>

Note: *Significant at 1 per cent level.  **Significant at 5 per cent level

Elasticity coefficients relate to cross-section of 14 major States. F.C.= Finance Commission.

Source: Estimated from the data taken from the Budget Documents of the State Governments.
Table 5-7: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statutory Transfers per capita (Rs.)</td>
<td>57.78</td>
<td>14.63</td>
</tr>
<tr>
<td>Grants for State Plan Schemes per capita (Rs.)</td>
<td>13.97</td>
<td>7.12</td>
</tr>
<tr>
<td>Discretionary Transfers per capita (Rs.)</td>
<td>18.61</td>
<td>6.08</td>
</tr>
<tr>
<td>Total Transfers per capita (Rs.)</td>
<td>90.36</td>
<td>20.82</td>
</tr>
<tr>
<td>State Domestic Product (Rs. 10,000,000)</td>
<td>995,800</td>
<td>549,135</td>
</tr>
<tr>
<td>State Domestic Product per capita (Rs.)</td>
<td>2027</td>
<td>724</td>
</tr>
<tr>
<td>Population (millions)</td>
<td>53.07</td>
<td>29.21</td>
</tr>
<tr>
<td>Proportion of MPs from state in ruling party/coalition</td>
<td>0.066</td>
<td>0.054</td>
</tr>
<tr>
<td>Match between central and state ruling parties</td>
<td>0.55</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Note: All financial variables are measured in 1981 Rupees

Correlations of Independent Variables

<table>
<thead>
<tr>
<th></th>
<th>SDP</th>
<th>SDP capita</th>
<th>Population</th>
<th>Match</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDP</td>
<td>0.1826</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDP per capita</td>
<td></td>
<td>0.7831</td>
<td>-0.3207</td>
<td>0.0999</td>
<td>0.4485</td>
</tr>
<tr>
<td>Population</td>
<td>0.1388</td>
<td>0.0654</td>
<td>0.0999</td>
<td>0.6210</td>
<td>0.4485</td>
</tr>
<tr>
<td>Match</td>
<td>0.4853</td>
<td>-0.2291</td>
<td>0.6210</td>
<td>0.4485</td>
<td></td>
</tr>
</tbody>
</table>
Table 5-8: Linear Specification Coefficients

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) Statutory Transfers per capita</th>
<th>(2) Statutory Transfers per capita</th>
<th>(3) Grants for State Plan Schemes per capita</th>
<th>(4) Discretionary Transfers per capita</th>
<th>(5) Total Transfers per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDP</td>
<td>-0.162E-04 (-1.262)</td>
<td>-0.173E-04 (-1.328)</td>
<td>0.328E-06 (0.036)</td>
<td>-0.912E-05 (-1.475)</td>
<td>-0.261E-04** (-2.151)</td>
</tr>
<tr>
<td>SDPPC</td>
<td>0.0142** (2.171)</td>
<td>0.0156** (2.363)</td>
<td>0.586E-04 (0.013)</td>
<td>0.499E-02 (1.592)</td>
<td>0.0206** (2.219)</td>
</tr>
<tr>
<td>POPN</td>
<td>1.224*** (2.919)</td>
<td>1.216*** (2.769)</td>
<td>0.295 (0.950)</td>
<td>0.812*** (3.888)</td>
<td>2.323*** (6.183)</td>
</tr>
<tr>
<td>MATCHLAG3</td>
<td>1.7344 (0.636)</td>
<td>4.784** (2.481)</td>
<td>0.110 (0.085)</td>
<td>6.628 (1.631)</td>
<td></td>
</tr>
<tr>
<td>PROPNLG3</td>
<td>52.684 (1.107)</td>
<td>4.687 (0.139)</td>
<td>5.040 (0.223)</td>
<td>62.440 (1.194)</td>
<td></td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.62681</td>
<td>0.62802</td>
<td>0.21306</td>
<td>0.51408</td>
<td>0.62325</td>
</tr>
</tbody>
</table>

Note: All financial variables are measured in 1981 Rupees

$t$-ratios in parentheses

*** significant at 1% level, ** significant at 5% level, * significant at 10% level (all two-sided)
### Table 5-9: Linear Specification Marginal Effects

<table>
<thead>
<tr>
<th>Impact Variable (Held Constant)</th>
<th>(2) Statutory Transfers per capita</th>
<th>(4) Discretionary Transfers per capita</th>
<th>(5) Total Transfers per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDP (SDPPC)</td>
<td>0.43E-04***</td>
<td>0.31E-04***</td>
<td>0.93E-04***</td>
</tr>
<tr>
<td>(SDPPC)</td>
<td>(3.14)</td>
<td>(4.78)</td>
<td>(6.52)</td>
</tr>
<tr>
<td>SDP (POPN)</td>
<td>0.12E-04</td>
<td>0.28E-06</td>
<td>0.11E-04</td>
</tr>
<tr>
<td>(POPN)</td>
<td>(1.48)</td>
<td>(0.07)</td>
<td>(1.26)</td>
</tr>
<tr>
<td>SDPPC (SDP)</td>
<td>-0.014</td>
<td>-0.015***</td>
<td>-0.039***</td>
</tr>
<tr>
<td>(SDP)</td>
<td>(-1.41)</td>
<td>(-3.14)</td>
<td>(-3.24)</td>
</tr>
<tr>
<td>SDPPC (POPN)</td>
<td>0.006</td>
<td>0.15E-03</td>
<td>0.006</td>
</tr>
<tr>
<td>(POPN)</td>
<td>(1.48)</td>
<td>(0.07)</td>
<td>(1.26)</td>
</tr>
<tr>
<td>POPN (SDP)</td>
<td>0.66*</td>
<td>0.64***</td>
<td>1.74***</td>
</tr>
<tr>
<td>(SDP)</td>
<td>(1.67)</td>
<td>(3.35)</td>
<td>(3.76)</td>
</tr>
<tr>
<td>POPN (SDPPC)</td>
<td>0.87***</td>
<td>0.63***</td>
<td>1.89***</td>
</tr>
<tr>
<td>(SDPPC)</td>
<td>(3.14)</td>
<td>(4.78)</td>
<td>(6.52)</td>
</tr>
</tbody>
</table>

T-ratios in parentheses

*** significant at 1% level, ** significant at 5% level, * significant at 10% level (all two-sided)
Table 5-10: Linear Specification Fixed Effects

<table>
<thead>
<tr>
<th>State</th>
<th>(2) Statutory Transfers per capita</th>
<th>(3) Grants for State Plan Schemes per capita</th>
<th>(4) Discretionary Transfers per capita</th>
<th>(5) Total Transfers per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andhra Pradesh</td>
<td>-32.282</td>
<td>-10.041</td>
<td>-31.045</td>
<td>-73.403</td>
</tr>
<tr>
<td>Bihar</td>
<td>-47.652</td>
<td>-18.333</td>
<td>-47.100</td>
<td>-113.122</td>
</tr>
<tr>
<td>Haryana</td>
<td>-15.283</td>
<td>7.830</td>
<td>-1.928</td>
<td>-9.403</td>
</tr>
<tr>
<td>Karnataka</td>
<td>-21.737</td>
<td>-6.311</td>
<td>-17.990</td>
<td>-46.066</td>
</tr>
<tr>
<td>Kerala</td>
<td>5.690</td>
<td>4.207</td>
<td>-11.362</td>
<td>-1.487</td>
</tr>
<tr>
<td>Madhya Pradesh</td>
<td>-34.247</td>
<td>-9.233</td>
<td>-32.175</td>
<td>-75.685</td>
</tr>
<tr>
<td>Maharashtra</td>
<td>-52.746</td>
<td>-17.923</td>
<td>-35.963</td>
<td>-106.667</td>
</tr>
<tr>
<td>Orissa</td>
<td>33.165</td>
<td>6.709</td>
<td>-2.121</td>
<td>37.737</td>
</tr>
<tr>
<td>Punjab</td>
<td>-19.024</td>
<td>3.608</td>
<td>-10.104</td>
<td>-25.547</td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>-18.331</td>
<td>-5.330</td>
<td>-26.070</td>
<td>-49.759</td>
</tr>
<tr>
<td>Uttar Pradesh</td>
<td>-96.628</td>
<td>-27.150</td>
<td>-77.766</td>
<td>-201.597</td>
</tr>
</tbody>
</table>
Table 5-11: LogLinear Specification Coefficients

<table>
<thead>
<tr>
<th>Variable</th>
<th>(6) Statutory Transfers per capita</th>
<th>(7) Grants for State Plan Schemes per capita</th>
<th>(8) Discretionary Transfers per capita</th>
<th>(9) Total Transfers per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNSDPPC</td>
<td>0.397* (1.868)</td>
<td>1.086 (1.492)</td>
<td>-0.299 (-0.923)</td>
<td>0.380** (1.974)</td>
</tr>
<tr>
<td>LNPOPN</td>
<td>0.661** (1.979)</td>
<td>-0.745 (-0.657)</td>
<td>2.704*** (4.868)</td>
<td>0.852*** (2.904)</td>
</tr>
<tr>
<td>MATCHLAG3</td>
<td>0.0093 (0.230)</td>
<td>0.253* (1.950)</td>
<td>-0.012 (-0.196)</td>
<td>0.054 (1.316)</td>
</tr>
<tr>
<td>PROPNLAG3</td>
<td>1.397** (2.428)</td>
<td>1.637 (1.222)</td>
<td>1.384 (1.353)</td>
<td>1.300** (2.451)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.60946</td>
<td>0.25657</td>
<td>0.49594</td>
<td>0.60212</td>
</tr>
<tr>
<td>R-squared</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: All financial variables are measured in 1981 Rupees
t-ratios in parentheses
*** significant at 1% level, ** significant at 5% level, * significant at 10 % level (all two-sided)
Table 5-12: LogLinear Specification Fixed Effects

<table>
<thead>
<tr>
<th>State</th>
<th>(6) Statutory Transfers per capita</th>
<th>(7) Grants for State Plan Schemes per capita</th>
<th>(8) Discretionary Transfers per capita</th>
<th>(9) Total Transfers per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andhra Pradesh</td>
<td>-1.760</td>
<td>-2.833</td>
<td>-6.129</td>
<td>-2.006</td>
</tr>
<tr>
<td>Gujarat</td>
<td>-1.885</td>
<td>-3.482</td>
<td>-5.021</td>
<td>-1.930</td>
</tr>
<tr>
<td>Haryana</td>
<td>-1.236</td>
<td>-4.267</td>
<td>-1.961</td>
<td>-1.040</td>
</tr>
<tr>
<td>Karnataka</td>
<td>-1.667</td>
<td>-3.450</td>
<td>-5.052</td>
<td>-1.822</td>
</tr>
<tr>
<td>Kerala</td>
<td>-1.121</td>
<td>-3.053</td>
<td>-4.134</td>
<td>-1.230</td>
</tr>
<tr>
<td>Madhya Pradesh</td>
<td>-1.765</td>
<td>-2.605</td>
<td>-6.256</td>
<td>-2.001</td>
</tr>
<tr>
<td>Maharashtra</td>
<td>-2.218</td>
<td>-3.629</td>
<td>-6.447</td>
<td>-2.497</td>
</tr>
<tr>
<td>Orissa</td>
<td>-0.748</td>
<td>-2.700</td>
<td>-3.851</td>
<td>-0.885</td>
</tr>
<tr>
<td>Punjab</td>
<td>-1.343</td>
<td>-4.541</td>
<td>-2.804</td>
<td>-1.320</td>
</tr>
<tr>
<td>Rajasthan</td>
<td>-1.425</td>
<td>-2.746</td>
<td>-4.667</td>
<td>-1.462</td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>-1.596</td>
<td>-2.947</td>
<td>-5.726</td>
<td>-1.839</td>
</tr>
<tr>
<td>Uttar Pradesh</td>
<td>-2.253</td>
<td>-2.009</td>
<td>-8.312</td>
<td>-2.625</td>
</tr>
<tr>
<td>West Bengal</td>
<td>-1.597</td>
<td>-2.733</td>
<td>-6.664</td>
<td>-1.954</td>
</tr>
<tr>
<td>Variable</td>
<td>(10) Statutory Transfers per capita</td>
<td>(11) Grants for State Plan Schemes per capita</td>
<td>(12) Discretionary Transfers per capita</td>
<td>(13) Total Transfers per capita</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td></td>
<td>(3.324)</td>
<td>(1.430)</td>
<td>(1.513)</td>
<td>(3.074)</td>
</tr>
<tr>
<td>LNSDPPC^2</td>
<td>-0.554**</td>
<td>-0.700</td>
<td>-0.315</td>
<td>-0.431**</td>
</tr>
<tr>
<td></td>
<td>(-2.351)</td>
<td>(-1.276)</td>
<td>(-0.952)</td>
<td>(-2.235)</td>
</tr>
<tr>
<td>LNPOPN</td>
<td>1.396</td>
<td>-7.460*</td>
<td>1.976</td>
<td>-0.343</td>
</tr>
<tr>
<td></td>
<td>(0.964)</td>
<td>(-1.855)</td>
<td>(0.830)</td>
<td>(-0.236)</td>
</tr>
<tr>
<td>LNPOPN^2</td>
<td>0.059</td>
<td>0.177</td>
<td>0.570</td>
<td>0.253</td>
</tr>
<tr>
<td></td>
<td>(0.178)</td>
<td>(0.184)</td>
<td>(1.264)</td>
<td>(0.960)</td>
</tr>
<tr>
<td>LNSDPPC*</td>
<td>-0.149</td>
<td>0.708</td>
<td>-0.497</td>
<td>-0.101</td>
</tr>
<tr>
<td>LNPOPN</td>
<td>(-0.369)</td>
<td>(0.578)</td>
<td>(-0.872)</td>
<td>(-0.295)</td>
</tr>
<tr>
<td>MATCHLAG3</td>
<td>-0.136E-02</td>
<td>0.239*</td>
<td>-0.363E-02</td>
<td>0.505E-01</td>
</tr>
<tr>
<td></td>
<td>(-0.034)</td>
<td>(1.959)</td>
<td>(-0.060)</td>
<td>(1.397)</td>
</tr>
<tr>
<td>PROPNALG3</td>
<td>1.373**</td>
<td>1.334</td>
<td>0.979</td>
<td>1.095**</td>
</tr>
<tr>
<td></td>
<td>(2.279)</td>
<td>(1.002)</td>
<td>(0.895)</td>
<td>(2.081)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.63575</td>
<td>0.26837</td>
<td>0.50006</td>
<td>0.63185</td>
</tr>
</tbody>
</table>

Note: All financial variables are measured in 1981 Rupees

T-ratios in parentheses

*** significant at 1% level, ** significant at 5% level, * significant at 10% level (all two-sided)
<table>
<thead>
<tr>
<th>State</th>
<th>(10) Statutory Transfers per capita</th>
<th>(11) Grants for State Plan Schemes per capita</th>
<th>(12) Discretionary Transfers per capita</th>
<th>(13) Total Transfers per capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haryana</td>
<td>-36.990</td>
<td>-22.149</td>
<td>-27.383</td>
<td>-25.769</td>
</tr>
<tr>
<td>Punjab</td>
<td>-37.006</td>
<td>-22.249</td>
<td>-27.924</td>
<td>-25.881</td>
</tr>
</tbody>
</table>
## Table 5A-1

TAXATION HEADS ASSIGNED TO THE UNION AND THE STATES IN THE CONSTITUTION  
(AS LISTED IN THE SEVENTH SCHEDULE OF THE CONSTITUTION)

<table>
<thead>
<tr>
<th>Entry in List I of the Seventh Schedule</th>
<th>Head</th>
<th>Entry in List II of the Seventh Schedule</th>
<th>Head</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>Taxes on income other than agricultural income.</td>
<td>45</td>
<td>Land revenue, including the assessment and collection of revenue, the maintenance of land records, survey for revenue purposes.</td>
</tr>
<tr>
<td>83</td>
<td>Duties of customs including export duties.</td>
<td>46</td>
<td>Taxes on agricultural income.</td>
</tr>
<tr>
<td>84</td>
<td>Duties of excise on tobacco and other goods manufactured or produced in India except- a. alcoholic liquors for human consumption; b. opium, Indian hemp and other narcotic drugs and narcotics; but including medicinal and toilet preparations containing alcohol or any substance included in sub-paragraph (b) of this entry.</td>
<td>47</td>
<td>Duties in respect of succession of agricultural land.</td>
</tr>
<tr>
<td>85</td>
<td>Corporation tax.</td>
<td>48</td>
<td>Estate duty in respect of agricultural land.</td>
</tr>
<tr>
<td>86</td>
<td>Taxes on the capital value of the assets, exclusive of agricultural land of individuals and companies; taxes on the capital of companies.</td>
<td>49</td>
<td>Taxes on lands and buildings.</td>
</tr>
<tr>
<td>87</td>
<td>Estate duty in respect of property other than agricultural land.</td>
<td>50</td>
<td>Taxes on mineral rights subject to any limitations imposed by Parliament by law relating to mineral development.</td>
</tr>
<tr>
<td>88</td>
<td>Duties in respect of succession to property other than agricultural land</td>
<td>51</td>
<td>Duties of excise on the following goods manufactured or produced in the State and countervailing duties at the same or lower rates on similar goods manufactured or produced elsewhere in India: a. alcohol liquors for human consumption; b. opium, Indian hemp and other narcotic drugs and narcotics; but not including medicinal and toilet preparations containing alcohol or any substance included in sub-paragraph (b) of this entry.</td>
</tr>
<tr>
<td>89</td>
<td>Terminal taxes on goods or passengers carried by railway, sea or air; taxes on railway fares and freights.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Taxes on the entry of goods into a local area for consumption, use or sale therein.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td><strong>90</strong></td>
<td>Taxes other than stamp duties on transactions in stock exchanges and future markets.</td>
<td><strong>53</strong></td>
<td>Taxes on the consumption or sale of electricity.</td>
</tr>
<tr>
<td><strong>91</strong></td>
<td>Rates of stamp duty in respect of bills of exchange cheques promissory notes, bills of lading, letters of credit, policies of insurance, transfer of shares, debentures, proxies and receipts.</td>
<td>@54</td>
<td>Taxes on the sale or purchase of goods other than newspapers, subject to the provisions of entry 92A of List I.</td>
</tr>
<tr>
<td><strong>92</strong></td>
<td>Taxes on the sale or purchase of newspapers and on advertisements published therein.</td>
<td><strong>55</strong></td>
<td>Taxes on advertisements other than advertisements published in the newspaper @@ and advertisements broadcast by radio or television.</td>
</tr>
<tr>
<td><em>92A</em></td>
<td>Taxes on the sale or purchase of goods other than newspapers, where such sale or purchase takes place in the course of inter-State trade or commerce.</td>
<td><strong>56</strong></td>
<td>Taxes on goods and passengers carried by road or on inland waterways.</td>
</tr>
<tr>
<td><strong>92B</strong></td>
<td>Taxes on the consignment of goods (whether the consignment is to the person making it or to any other person), where such consignment takes place in the course of inter-State trade or commerce.</td>
<td><strong>57</strong></td>
<td>Taxes on vehicles, whether mechanically propelled or not, suitable for use on roads including tramcars, subject to the provision of entry 35 of List III.</td>
</tr>
<tr>
<td><strong>97</strong></td>
<td>Any other matter not enumerated in List II or List III including any tax not mentioned in either or both the Lists.</td>
<td><strong>58</strong></td>
<td>Taxes on animals and boats.</td>
</tr>
<tr>
<td></td>
<td><strong>59</strong></td>
<td>Tolls.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>60</strong></td>
<td>Taxes on professions, trades, callings and employments.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>61</strong></td>
<td>Capitation taxes.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>62</strong></td>
<td>Taxes on luxuries, including taxes on entertainments, amusements, betting and gambling.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>63</strong></td>
<td>Rates of stamp duty in respect of documents other than those specified in the provision of List I with regard to rates of stamp duty.</td>
<td></td>
</tr>
</tbody>
</table>

* Ins. by the Constitution (Sixth Amendment) Act, 1956 s.2  
** Ins. by the Constitution (Forty-sixth Amendment) Act, 1982, s.5  
@ Sub. by the Constitution (sixth Amendment) Act 1956, s.2 for entry 54  
@@ Ins. by the Constitution (Forth-second Amendment) Act, 1975, s.57 (w.e.f. 31.1.1977)  

118
### Table 5A-2

**Distribution of the States' Share in the Net Proceeds of Non-corporate Income-tax**

<table>
<thead>
<tr>
<th>Finance Commissions</th>
<th>Net Proceeds distributed to the States</th>
<th>Criteria for Distribution</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contribution</td>
<td>Population</td>
<td>Per capita SDP</td>
</tr>
<tr>
<td>First</td>
<td>50</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>Second</td>
<td>60</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Third</td>
<td>60.67</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>Fourth</td>
<td>75</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>Fifth</td>
<td>75</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Sixth</td>
<td>80</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Seventh</td>
<td>85</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Eighth</td>
<td>85</td>
<td>10</td>
<td>22.5</td>
</tr>
<tr>
<td>Ninth (First Report)</td>
<td>85</td>
<td>10</td>
<td>22.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ninth (Second Report)</td>
<td>85</td>
<td>10</td>
<td>22.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenth</td>
<td>77.5</td>
<td>-</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* According to "distance" formula - see notes under Table 5A-3.
** According to "inverse" formula - see notes under Table 5A-3.
@ The variables included are (i) the population of scheduled castes and tribes; and (ii) number of agricultural laborers. Equal weights are assigned to the two factors.
<table>
<thead>
<tr>
<th>Finance Commissions</th>
<th>Coverage</th>
<th>States' share (per cent)</th>
<th>Proportion of population of the State to the total population of all States</th>
<th>Per capita income</th>
<th>Economic or social backwardness</th>
<th>Other criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 First</td>
<td>Three commodities: tobacco, matches and vegetable products</td>
<td>40</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 Second</td>
<td>Eight commodities:</td>
<td>25</td>
<td>90</td>
<td>-</td>
<td>-</td>
<td>10 per cent for adjustment</td>
</tr>
<tr>
<td>3 Third</td>
<td>All commodities yielding more than Rs 5.9 million in 1960-61 (about 35)</td>
<td>20</td>
<td>Mainly population basis along with relative financial weakness and economic backwardness as other factors.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4 Fourth</td>
<td>All commodities excluding regulatory duties, special excises and earmarked cesses</td>
<td>20</td>
<td>80</td>
<td>-</td>
<td>20 Backwardness as indicated by seven factors: i) per capita agricultural production; ii) per capita manufacturing value added; iii) percentage of workers to total population, iv) percentage of enrollment in class 1 to 5 to the population in the age group 6-11, v) population per hospital bed, vi) percentage of rural population, vii) percentage of scheduled caste population</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Fifth</strong></td>
<td>All types of union excise duties (for the first 3 years (1969-72), Regulatory duties and earmarked cesses are excluded</td>
<td>20</td>
<td>60</td>
<td>13.3 Distributed among only those States whose per capita SDP was below all States average, in proportion to the shortfall of the State's per capita SDP from all State average multiplied by the population of the State</td>
<td>6.7 According to an integrated index of backwardness measured by: i) scheduled caste population, ii) number of factory workers per lakh of population, iii) net irrigated area per cultivator, iv) length of railways and surfaced roads per square kilometre area, v) enrollment ratio of school going age children; and number of hospitals beds per thousand person</td>
<td></td>
</tr>
<tr>
<td><strong>Sixth</strong></td>
<td>For 1974-75 and 1975-76 all items except auxiliary duties of excise and cesses levied under special acts and earmarked for special purposes</td>
<td>20</td>
<td>75</td>
<td>25 According to the &quot;distance&quot; formula</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td><strong>Seventh</strong></td>
<td>All items excluding duty on the generation of electricity</td>
<td>45</td>
<td>25</td>
<td>25 Inverse* of per capita SDP formula</td>
<td>25 Percentage of poor</td>
<td>25 According to a formula equalizing revenue capacity computed by regressing States' per capita revenue on per capita SDP and substituting the actual values of per capita SDP in the equation.</td>
</tr>
<tr>
<td><strong>Eighth</strong></td>
<td>Net proceeds: excluding cesses levied under special Acts and earmarked for special purposes</td>
<td>45</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>(5 per cent to deficit States) in proportion to the deficit of a State to the total deficit of the State in that year</td>
</tr>
</tbody>
</table>

121
<table>
<thead>
<tr>
<th>Ninth First Report (1989-90)</th>
<th>Net proceeds excluding cesses levied under special acts and earmarked cesses</th>
<th>45 (40 per cent to all States and 5 per cent to the States having post-devolution deficits)</th>
<th>25</th>
<th>50</th>
<th>12.5 Percentage of people below poverty line</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ninth Second Report (1990-95)</td>
<td>Net proceeds excluding cesses levied under Special Acts and earmarked cesses</td>
<td>45 (40 per cent for all States. 5 per cent for the States with post-devolution deficits.)</td>
<td>25</td>
<td>12.5 Inverse* formula</td>
<td>12.5 Index of backwardness computed with equal weights assigned to population of scheduled castes and tribes and number of agricultural laborers</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33.5 Distance formula**</td>
<td></td>
</tr>
<tr>
<td>Tenth (1995-2000)</td>
<td>Net proceeds excluding cesses levied under special Acts and earmarked cesses</td>
<td>47.5 (40 per cent to all the States and 7.5 per cent to the States having post-devolution deficits)</td>
<td>20</td>
<td>60</td>
<td>5 (index of Infrastructure)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Distance formula</td>
<td>Tax effort as measured by the ratio of per capita tax revenue to the square of per capita SDP in the state scaled by population.</td>
</tr>
</tbody>
</table>

*Inverse formula = \( \frac{P_i}{Y_i} \cdot \frac{\sum P_i}{Y_i} \)

**Distance formula = \( \frac{(Y_h - Y_i)P_i}{(Y_h - Y_i)P_i} \)

where \( Y_i \) and \( Y_h \) represent per capita SDP of the \( i^{th} \) and the highest per capita SDP State, \( P_i \) - the population of the \( i^{th} \) State, \( (Y_h - Y_i) \) for the 'h' State is taken to be the distance between the highest and the next highest per capita SDP.
Table 5A-4
Formula for Distributing State Plan Assistance*

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Share in central plan assistance (per cent)</th>
<th>Share of grants and loans</th>
<th>Criteria for distribution in non-special category States</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Special category States (10)</td>
<td>30</td>
<td>90:10</td>
<td></td>
</tr>
<tr>
<td>B. Non-special category States (15)</td>
<td>70</td>
<td>30:70</td>
<td></td>
</tr>
<tr>
<td>(i) Population (1971)</td>
<td></td>
<td>60.0</td>
<td></td>
</tr>
<tr>
<td>(ii) Per capita income, of which</td>
<td></td>
<td>25.0</td>
<td></td>
</tr>
<tr>
<td>(a) According to the 'deviation' method covering only the States with per capita income below the national average</td>
<td></td>
<td>20.0</td>
<td></td>
</tr>
<tr>
<td>(b) According to the 'distance' method covering all the fifteen States</td>
<td></td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>(iii) Fiscal performance, of which</td>
<td></td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>(a) Tax effort</td>
<td></td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>(b) Fiscal management</td>
<td></td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>(c) National objectives</td>
<td></td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>(iv) Special problems</td>
<td></td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Note: 1. The formula as revised in December, 1991.
2. Fiscal management is assessed as the difference between States' own total plan resources estimated at the time of finalizing annual plan and their actual performance, considering latest five years.
3. Under the criterion of the performance in respect of certain programs of national priorities the approved formula covers four objectives, viz. (i) population control, (ii) elimination of illiteracy, (iii) on-time completion of externally aided projects, and (iv) success in land reforms.
Chapter 6

Corruption and the Organization of Tax Administration: Non-Smithian Benefits from Specialization

John McLaren

This note attempts to formalize and study an idea that has emerged as an important one in the reform of tax administration: The idea of functional specialization as a hedge against corruption.

The idea simply put is as follows. There are many tasks that go into enforcing a tax code. Potential voters need to be identified and registered. A sample of them need to be selected for audit; audits involve a review of documents submitted by the taxpayer, and perhaps a visit to the home or place of business; documents may be sought from third parties; in some cases co-workers may be interviewed. Once information has been assembled, it needs to be evaluated and a judgement made, appeals may need to be dealt with, and enforcement actions need to be followed through. In short, the simple idea of "enforcing the tax code" requires a large number of different, coordinated activities.

There are roughly two ways of organizing these activities. One may assign a district director with a staff and charge that director with all aspects of enforcement for the district. This is, roughly, the Indian system (Das-Gupta and Mookherjee (1998, chapter 6)). This might be called "geographic specialization" or "decentralization," but here we will call it a "generalist system." On the other hand, for each task one may set up a corps that specializes in that task, covering a broad geographic area, and require those different corps to work together in each district. This is what is meant by "functional specialization" (we will say "specialization" for short).

The trend in the reform of tax administration is toward greater functional specialization in a large number of countries, because of perceived advantages in direct efficiency (of the type that Adam Smith famously described in his pin factory) and also because of advantages in "depersonalizing" tax enforcement and thus reducing corruption. However, it can also come at a cost, particularly in many low-income countries, because of the need for the different functional groups to be well coordinated and able to share and analyze information together. Thus, a poor information infrastructure including weak computerization can lead to large inefficiencies from

---

* I am grateful to Margaret Madajewicz, Dilip Mookherjee, and to conference participants for comments. All errors and deficiencies are my own.
functional specialization, and this appears to be a binding constraint in the Indian case.\footnote{See Das-Gupta and Mookherjee 1998, pp. 259-60. The rise of functional specialization is a major theme of that work, and it recurs in the discussion of several country case studies.}

A point made in this note is that even in the presence of such direct inefficiencies, functional specialization can be a profitable policy, because of what it can do to incentives for corruption. Essentially, what it does is to move from a system in which an official (or team) handles all of the administration of a small number of taxpayers, to a situation in which the official (or team) handles a portion of the administration of a large number of taxpayers. In consequence, each taxpayer now deals with a large number of officials (or teams), and in order for the taxpayer to bribe his or her way out of a tax liability it would be necessary to bribe a large number of people. This is a more expensive kind of bribery; the rents will need to be spread over a larger group of people; and in consequence it is less likely to be a viable strategy.\footnote{The argument has some loose parallels in industrial organization theory. There is a distant family relationship between this and the well-known principle that collusion is harder to sustain in a large oligopoly than in a small one (see Shapiro (1989)). In both cases, the spreading of rents over a large number of agents is important.}

It should be pointed out that this note is not the first appearance of the idea in a formal model; a version has appeared in a model of bureaucratic corruption in the important survey by Rose-Ackerman (1978, p. 183). However, there it comes up essentially in passing; to my knowledge the argument has never been given a complete theoretical treatment, including an analysis of the conditions under which functional specialization is optimal. The following sections attempt to do this.

\textit{I. The Model}

Consider a model of an inland revenue agency along the following lines. For the most part, this is a familiar ‘principal-agent-client’ model of bureaucratic corruption, with the specialization issue being the only new element (see Bardhan 1997 or McLaren 1996 for surveys). The structure is mapped out in Figure 6-1. Suppose that there are two functions that may be performed by an auditor, labelled A and B. Any auditor can perform either function; both take the same amount of time; and both are applicable to any taxpayer. In the time available, each auditor can perform A audits on n taxpayers, B audits on n taxpayers, or any combination of A and B audits such that their numbers add up to n, where n is an even number. Further, the government has hired enough auditors that there are \(n/2\) taxpayers per auditor.

There are two types of taxpayer: High-income and low-income. High-income taxpayers owe \(\tau_H\) in tax, and low-income taxpayers owe \(\tau_L\) in tax, with \(\Delta \tau = \tau_H - \tau_L > 0\). Low-income taxpayers make up a fraction \(\lambda\) of the total (think of \(\lambda\) as standing for ‘low-income’). Each taxpayer must file a tax return stating his or her income, and understands that if the Inland Revenue Service finds that she has claimed falsely to have a low income, she will be assessed a fine amounting to \(F_T\).

A taxpayer who files dishonestly faces some probability that that deception will be found out by an auditor. There are two possible ways in which this can happen, depending on the structure
Figure 6-1. Decision Tree for the Case of a High-Income Taxpayer

- Truthful return.
- Dishonest return.
  - Auditor finds discrepancy.
  - Auditor finds no discrepancy.
- Collusion.
  - Superiors detect collusion.
  - Superiors do not detect collusion.

Taxpayer's decision.

Superiors

Auditor finds

Report
discrepancy
to superiors.

(1-q)

(1-φ)

Dishonest

φ

return.
chosen by the top administration. The first system is a ‘generalist’ system, in which each auditor is responsible for auditing a certain number of taxpayers, performing both A and B functions. Under this system, the probability that a dishonest taxpayer will be found out is given by the parameter \( \phi \) (think of \( \phi \) as standing for ‘found out’). The second is a ‘specialist’ system, in which each auditor is responsible either for performing A audits only on a given set of taxpayers, or for performing B audits only on a given set of taxpayers. Under this system, the probability that a dishonest taxpayer is found out by his or her A auditor or by the B auditor but not both is given by \( \phi' \). The probability that the deception is found out by both auditors is given by \( \phi'' \). This could be either by both independently realizing the same discrepancy in different aspects of the records, or by communicating and finding that the records A is inspecting are inconsistent with the records that B is inspecting. Either way, assume that when both inspectors find that the return in fraudulent, that becomes common knowledge between them and the taxpayer. In sum, the probability of evading detection is given by \( 1 - \phi' - \phi'' \).

It is possible that either \( \phi < \phi' + \phi'' \) or that \( \phi > \phi' + \phi'' \). In the first case there are technological advantages to specialization, presumably from learning effects or indivisibilities in the time-costs of each activity. We can call this the ‘Smithian’ case. In the second case, communications difficulties, inadequate record-keeping or computerization make it difficult for the two auditors working independently to spot an inconsistency that would have been noticed by a single auditor working alone to inspect all of the records of a given taxpayer. This is the potential drawback to functional specialization in a tax office with poor informational infrastructure, as emphasized by Das-Gupta and Mookherjee (1998, pp. 259-60). It will become clear that this is the interesting case, where there is a trade-off between the technological disadvantage of functional specialization and its advantages for incentives, and we will therefore assume it throughout. Call this case the ‘anti-Smithian’ case.

The auditors all are paid the same salary, which is a key decision variable for the government and whose value is denoted by \( \omega \). If an auditor finds that a return is fraudulent, he has two options. He can report the infraction to his superiors, in which case the taxpayer will be fined; or he can offer to keep silent in return for a bribe. Call this second case ‘collusion.’ Of course, in the specialist system, in the event that both auditors are aware of the deception, collusion will require three parties to agree.

If the taxpayer and auditor(s) do collude, they will be found out by the auditor’s superiors with probability \( 1 - \phi \). In that case, the taxpayer will pay the additional taxes owed plus the fine, and the auditor will be fired and lose his wage, and in addition will need to pay a fine of \( F_A \). (The auditor’s opportunity income has been normalized to zero.) Assume that there are no rewards for reporting evaded taxation and that the legal system works well enough that extortion is not an issue (that is, low-income taxpayers will not need to pay a bribe to prevent being declared high-income by the auditor). A full analysis would both allow for rewards (as in, for example, Mookherjee and P’ng (1995)) and deal with the problem of extortion directly (as in Hindriks, Keen, and Muthoo (1999), but this is beyond the purpose of this note.

All parties are risk-neutral and without ethical compunction, so that they maximize expected net income. Bargaining is governed by the Nash solution, generalized as necessary to three agents.
In addition, assume that the fines are capped by law or by liquidity constraints, and the government has already set them at their highest possible level.

II. Efficiency wages

In the case of a taxpayer who has been caught by one auditor, collusion will occur if the joint surplus from collusion exceeds the joint surplus from obedience to the Revenue Office. This holds if:

\[ q \omega + (1 - q)(- F_A - F_T - \Delta \tau) > \omega - F_T - \Delta \tau. \]

The left-hand side is the sum of the two payoffs from collusion (if they are not caught, the auditor gets a wage and nothing else happens, but if they are caught, the wage disappears and the taxpayer must pay taxes plus fines). The right-hand side corresponds to obedience. Thus, in order for collusion to be prevented, the remuneration to auditors must satisfy:

\[ \omega \geq \omega^e = (q/(1 - q))(\Delta \tau + F_T) - F_A. \]

Call this the ‘bilateral efficiency wage.’ If the wage is below this, the two will split the surplus by a bilateral bribe equal to:

\[ b = \frac{1}{2}[q(\Delta \tau + F_T) + (1 - q)(\omega + F_A)]. \]

In parallel, in the case of a taxpayer who has been caught by two auditors, collusion will occur if:

\[ 2q \omega + (1 - q)(- 2F_A - 2F_T - \Delta \tau) > 2 \omega - 2F_T - 2\Delta \tau. \]

The only difference with the corresponding condition for the bilateral case is that now two salaries and two auditor fines are at stake. The same level of taxes and taxpayer punishments are in place. A consequence is that the condition for collusion to be prevented is:

\[ \omega \geq \omega^e = \frac{1}{2} (q/(1 - q))(\Delta \tau + F_T) - F_A. \]

Call this the ‘trilateral efficiency wage.’ This is substantially lower than the bilateral, and if \( F_A = 0 \) it is exactly half. The reason is that the risk of the crime is the same per auditor, but with a larger number of conspirators the potential illicit profit per conspirator is now smaller. This is the heart of the incentive argument for specialization: It raises the number of people involved in each taxpayer case (without requiring any new staff), and thus makes it more difficult to conspire to conceal delinquent taxpayers.

The basic argument is illustrated by Figures 6-2 and 6-3. Figure 6-2 shows the organizational structure under generalist auditing. In this case, each auditor is responsible for all auditing of one taxpayer (here, \( n = 2 \)). In order for Taxpayer 4, for example, to buy her way out of trouble given that
Figure 6-2. The Auditing Structure with Generalist Auditors

A bribe circle.
Figure 6-3. The Auditing Structure with Specialist Auditors

Auditor 1

Auditor 2

Auditor 3

Auditor 4

Taxpayer 1

Taxpayer 2

Taxpayer 3

Taxpayer 4

A bribe circle.
her fraudulent return has been identified, she will need to collude with exactly one auditor, namely Auditor 4. This is illustrated by the bribe circle in the figure. By contrast, Figure 6-3 shows the specialist case. Auditors 1 and 3 are A-auditors. Each does the same amount of work as in Figure 6-2, but now does half the job of auditing two taxpayers, instead of the whole job for one. Similarly, Auditors 2 and 4 are B-auditors. Now, in the event that Taxpayer 4 has understated her income and has been detected by both of her auditors, in order for her to buy her way out of trouble she must collude with two auditors, which is more difficult and expensive and may not be worth it. This is illustrated by the larger bribe circle encompassing all necessary participants.

III. The Taxpayer's Disclosure Decision

At the stage of filing a return, a high-income taxpayer must choose between honesty and dishonesty. Assume for the moment a generalist system. There are two important regions of the parameter space to consider. In the first, the taxpayer would file a false return, even assuming that the auditor would never consider taking a bribe. Call this Case I. The second is the region in which the taxpayer would file a truthful return under that assumption. Call that situation Case II.

Assume Case I for the moment, because it is easier to deal with. The condition for this is:

$$\Delta T > \varphi F_T > (\varphi' + \varphi')F_T.$$  

Under Case I, the only revenues the government will receive will be \(r_L\) from each taxpayer, plus \(\Delta T + F_T\) from high-income audited taxpayers in the event that the government pays at least \(\omega^e\) to its auditors. A full analysis would need to consider the possibility that the optimal salary could be \(\omega^c\), zero (the opportunity income), or negative (it could be optimal to sell the jobs, thus implicitly collecting taxes through the price the auditors are willing to pay for their bribery rents -- an example of 'capitulation wages' (Besley and McLaren (1993)). However, for the present purpose, it is enough to focus on the case in which the efficiency wage is optimal (both under the specialist system and the generalist system). In this case, the government collects

$$[r_L + (1 - \lambda)\varphi(\Delta T + F_T)]$$  

per taxpayer, and spends \(\omega^e\) per auditor. Now consider the specialist system. If the government pays auditors \(\omega^e\), it will deter both bilateral and trilateral bribery and reap

$$[r_L + (1 - \lambda)(\varphi' + \varphi')(\Delta T + F_T)]$$  

per taxpayer. Under the anti-Smithian case this cannot be optimal, since it yields lower revenue than the generalist system at the same cost.

On the other hand, if the auditors are paid \(\omega^e\), trilateral but not bilateral collusion will be deterred. The result will be revenues of

$$[r_L + (1 - \lambda)\varphi'(\Delta T + F_T) + (1 - \lambda)\varphi(1 - q)(\Delta T + F_T + F_A)],$$  

(1)
per taxpayer, and costs of

\[1 - (1 - \lambda)\phi (1 - q)] \omega^e \quad (2)\]

per auditor. We know from the definition of \( \omega^e \) that

\[(1 - q)(\Delta \tau + F_T + F_A) = \Delta \tau + F_T - (1 - q)\omega^e.\]

Using this with (1) and subtracting (2) yields net revenue of:

\[\tau_e + (1 - \lambda)\phi (\gamma (\Delta \tau + F_T)) - [\omega^e + (1 - \lambda)\phi (1 - q)(\omega^e - \omega^e)].\]

Comparing this with net revenue under the generalist system, the first term in square brackets is smaller than the generalist revenue under our anti-Smithian assumption, and the second term in square brackets is smaller than the generalist wage cost because the trilateral efficiency wage is lower than the bilateral. This is the tradeoff discussed by Das-Gupta and Mookherjee: the first effect is the incomplete coordination of information across tax officials, and the second effect is essentially the benefit of the 'depersonalization' effect that makes corruption easier to deter.

Clearly, even in this anti-Smithian case, if \( \phi^e \) is not too far behind \( \phi \), specialization will be a profitable strategy. In particular, if \( \phi^e \mapsto \phi \) (so that \( \phi^e \mapsto 0 \)), clearly gross revenue under the specialist system converges to gross revenue under the generalist system, but the wage bill will remain substantially below. Thus, for high-\( \phi^e \) tax bureaucracies the revenue service should be run along lines of functional specialization and for low-\( \phi^e \) bureaucracies a generalist system will work better.

Briefly, Case II is more involved. It is easy to see that in Case II, paying the efficiency wage with generalists will never be optimal, because even if the auditors' pay is just below \( \omega^e \) and so auditors will be willing to accept bribes, taxpayers will file honestly. There will exist a level of auditor salary, say, \( \omega^T < \omega^e \), such that if auditors are paid less than \( \omega^T \), then the taxpayers will file dishonestly. Typically, this \( \omega^T \) will then be the optimum, and the analysis rests on a comparison of \( \omega^T \) under the two forms of organization. A similar trade-off applies.

Two final comments are in order. First, it is possible that some sociological complications could qualify this story. In the real world, it appears to be the case that a culture of corruption can be perpetuated by a social bond that develops between members of a team. An example that comes to mind is corruption in police departments, in which rookies can be introduced to local mores by their experienced partners, working together in the field away from supervision. Conceivably, some recruits who would never have considered taking bribes working on their own could be socialized into it by the incumbent members, if corruption is currently tolerated in the organization. If such

\[3\text{If the auditor's salary is slightly below } \omega^e, \text{ then the bargaining surplus between the auditor and an evading taxpayer whom she has caught is just slightly positive, which means that the evading taxpayer colluding with the auditor will be just slightly better off than an evading taxpayer who is simply turned in. (In other words, the bribe required to be paid is high enough that it is barely worth it.) Since the taxpayer was strictly willing to file an honest return when the auditor was being paid } \omega^e \text{ and thus not taking bribes, then he is still willing to do so.}\]
effects are possible, they could conceivably be a disadvantage to functional specialization, since such specialization certainly would require the staff to work in teams. Exploration of this possibility is beyond the scope of this note, however.4

Second, where specialization is optimal, it benefits the government by allowing it to lower auditor remuneration drastically without creating incentives for bribery. Clearly, although the government is better off under this outcome, the auditors are much worse off. This has implications for political economy. Surely auditors would have a strong incentive to resist specialization, and to resist the conditions under which it is possible. Indeed, they would rationally resist the computerization and improvements in information that (through a rise in $\phi$) make specialization desirable. Whether or not there is evidence of this in practice may be an interesting question to explore.

References


4I am grateful to participants in the conference discussion for raising this possibility.
Chapter 7

Incentive Effects of Performance-Based Rewards in Tax Administration*

Emilson C. D. Silva

Abstract. This paper illustrates some of the positive incentive effects of rewarding tax performance outcome indicators. By rewarding total collections, the government induces the autonomous tax administration to exert effort in improving the quality of tax enforcement. Improvements in quality effort, in turn, may increase the efficiency of both auditing of underreporters and inspecting of nonfilers. Although audits of underreporters and inspections of nonfilers are substitutes for each other, at least one of these enforcement quantities is necessarily positively related to enforcement quality. We also find that, under some reasonable circumstances, tax evasion falls with either an improvement in the computerization of the tax administration or an increase the highest income tax rate.

I. Introduction

Several developing countries have recently attempted to reform their tax collection agencies (see Bear, Castro and Vehorn 1999; Das-Gupta and Mookherjee 1998; Kahn, Silva and Ziliak 2001; Mookherjee 1997; Silvani and Baer 1997; and Vazquez-Caro, Reid and Bird 1992). Most reforms introduced or improved performance-related bonuses. In Brazil, for example, a tax collection reform introduced in 1988 provided monetary compensation to tax collectors based on their individual and group performance in finding and collecting taxes from tax evaders. Performance rewards were directly related to the amount of fines collected. Kahn, Silva and Ziliak (2001) find that the Brazilian bonus scheme led to a significant improvement in tax enforcement performance. Mexico also reformed its tax collection agency in 1988. According to Mookherjee (1997), the new Mexican system involves a bonus consisting of about 60 percent of additional collections within each office.

* This paper has been greatly improved following comments by John McLaren, Dilip Mookherjee and other participants in the Columbia-World Bank taxation conference on February, 18-19, 2000.
The implementation of a bonus system, however, is not per se sufficient for the success of the reform.\footnote{See, e.g., Baer, Castro and Vehorn 1999, Crotty 1998, Das-Gupta 1999 and Gill 1998 for a comprehensive list of measures and strategies that, if effectively implemented, may improve tax administration performance.} Das-Gupta and Mookherjee (1998), for example, show that tax enforcement in India has not been very effective in spite the existence of performance-based pay. It appears that the overall incentive effects of bonuses have been weak because both the likelihood of receiving a bonus and the size of the bonus are influenced by many factors beyond the control of a tax official.

The Indian case provides a good example of the many difficulties that there exist in improving tax administration performance with bonuses. Tax administrations are multitask institutions. Many of the tasks performed by tax administrations are not observable or not verifiable by the government, responsible for the payment of bonuses. Consider, for example, tax enforcement quality and field inspections of nonfilers. It is hard for the government to observe the degree of accuracy or diligence employed by a tax collector in the examination of returns. It is also very difficult for the government to monitor the performance of tax collectors when they conduct field inspections. This informational asymmetry implies that the government faces a moral hazard problem vis-a-vis the tax administration and that only verifiable outcomes can be rewarded. As Holmstrom and Milgrom (1991) have pointed out in more general contexts, the tax administration may have incentives to concentrate its efforts on verifiable margins of performance in lieu of unverifiable margins, such as quality and field inspections.

We show that the potential perverse effects associated with the implementation of tax performance bonuses – for example, low tax enforcement quality and low field inspection effort – need not to occur. If the government rewards the tax administration for its performance in aggregate outcomes, such as collection of taxes and fines, the tax administration may respond by improving tax enforcement quality and quantity measures. The underlying rationale for this behavior is quite straightforward. Performance rewards provide the tax administration with incentives to improve its effectiveness in the collection of overdue taxes and fines. The effectiveness of the tax administration is positively associated with both quantity and quality measures of tax enforcement. Indeed, improvements in quality measures may be necessary for improvements in the effectiveness of auditing of underreporters and inspecting of nonfilers. Consider, for example, an economy characterized by a low level of development and widespread tax evasion. It may initially be easy to catch tax evaders without raising tax enforcement quality because there is a large pool of tax evaders. Eventually, however, as the easy targets are gone, effective collection of overdue tax and fines will necessarily depend on the level of effort exerted by the tax administration in quality activities. By rewarding the tax administration for its performance in the collection of taxes and fines, the government may induce the tax administration to work harder in all enforcement activities, and subsequently raise the additional revenues needed to reach the government's targeted level of overall tax revenue.

Our economy consists of two sectors, black and white. Agents in the black sector may evade taxes by not filing at all or by underreporting income. Agents in the white sector, on the other hand, are assumed to always file tax returns. However, they may evade taxes by underreporting income. Our crucial finding is that, under some reasonable circumstances, the tax administration exerts effort in auditing of underreporters and inspecting of nonfilers at levels that
are positively related to the amount of effort it exerts in quality. In other words, auditing effort and inspecting effort may both be strategic complements to quality effort.

Given the fact that black markets are widespread in developing economies (see Goswami, Sanyal and Gang 1991 and McLaren 1996), we believe that our model may be quite useful in understanding how developing economies may effectively lower tax evasion. We find that improvements in computerization and increases in the highest tax rate may both motivate the tax administration to work harder in all of its enforcement activities, and an improvement in tax enforcement effectiveness may lower tax evasion because it raises the expected cost associated with tax evasion.

II. The Model

Consider an economy with two sectors, white (W) and black (B). The populations of agents in black and white sectors are \( P^b \) and \( P^w \), respectively. There are three types of income earners, high (H), medium (M), and low (L). The populations of income earners are \( P_H \), \( P_M \) and \( P_L \). We assume that \( P^b = P_L + P_H \) and \( P^w = P_M + P_H \). In words, agents in the black sector earn either low or high income. Agents in the white sector earn either medium or high income. Hence, \( P_H = P_H + P_M \).

In order to finance its expenditures, the government levies a tax on income. Before tax, the income levels are \( Y_H > Y_M > Y_L \). We assume that the income tax rates satisfy \( t_H > t_M > t_L \). We also assume that: (i) \( Y_L = 0 \); and (ii) \( t_L = 0 \). Low income earners ("low types") earn the subsistence level of income. Assumption (i) normalizes this income level to zero. Assumption (ii) tells us that low types do not pay taxes because they earn the subsistence level of income.

The government knows the population sizes and the levels of income earned by "whites" and "blacks," but it cannot immediately distinguish individuals according to their income types. To obtain information about individual incomes, the government relies on individual income reports and on the enforcement effort of its Tax Administration (TA). To simplify the analysis, we assume that whites must pre-register with the TA and that the TA can verify without cost whether or not whites filed tax returns. Our objective is to reserve the option of filing or not filing to earners of high black income ("high blacks").

Earners of white income have two options open to them: (a) report \( Y_M \); or (b) report \( Y_H \). High blacks, on the other hand, have three options open to them: (a) report \( Y_M \); (b) report \( Y_H \); or (c) do not file a tax return. We refer to high blacks that do not file as "high nonfilers" (\( N_n \)). Since low blacks do not file tax returns, the total population of nonfilers, \( N \), is \( N = P_L + N_n \).

The TA may monitor income earners in two possible ways: (i) by auditing returns or (ii) by conducting field inspections of nonfilers. We refer to audits of returns and field inspections of nonfilers simply as "audits" (A) and "inspections" (I), respectively. The TA may audit taxpayers who report medium income or audit taxpayers who report high income. Let \( A_M \) and \( A_H \) denote the numbers of medium-income and high-income audits. For simplicity, we assume that \( A_M = \alpha_M e_M, \alpha_M > 0 \), and \( A_H = \alpha_H e_H, \alpha_H > 0 \), where \( e_M \) and \( e_H \) represent the effort levels exerted by the
TA in auditing activities. Further, we assume that the number of inspections is also a linear function of the TA’s effort level exerted in this activity, \( I = \alpha_i e_i, \alpha_i > 0 \).

In addition to audits and inspections, the effectiveness of tax enforcement depends on the TA’s quality performance. Indeed, in the analysis below, quality performance is a necessary condition for the effectiveness of tax enforcement. The quality level, \( Q \), is a linear function of the effort exerted by the TA in this activity, \( Q = \alpha_o e_q, \alpha_o > 0 \).

Let \( E = e_M + e_H + e_q + e_o \). Hence, \( E \) denotes the total level of effort exerted by the TA. Let \( \theta \) denote the TA’s “degree of computerization.” We assume that \( 0 < \theta < 1 \). To keep things simple, we define the TA’s total operational cost function as follows:

\[
\Psi(e_M, e_H, e_i, e_o; \theta) = (1 - \theta)(e_M + e_H + e_i + e_o).
\]

It is important to note that the higher the TA’s degree of computerization, the lower its total and marginal costs.

Since the TA may allocate some of its effort away from quality activities toward audits and inspections, we define the TA’s “quality rate” as follows:

\[
q(e_q, E) = \frac{\alpha_o e_q}{\alpha_o E} = \frac{e_q}{E}.
\]

As we discuss below, we postulate that the likelihood that a tax evader faces punishment for tax evasion increases with the TA’s quality rate.

Let \( R_h \) and \( R_m \) denote the numbers of high and medium tax filers. Then, \( R_h = R^H_h + R^V_h \) and \( R_m = R^H_m + R^V_m \). The probabilities that high and medium tax filers are audited are assumed to be \( \left( \frac{A_h}{R_h} \right) \) and \( \left( \frac{A_m}{R_m} \right) \), respectively. Similarly, the probability that a nonfiler is inspected is assumed to be \( \left( \frac{1}{N} \right) \). These assumptions imply that the TA is unable to distinguish tax law offenders from abiders prior to auditing or inspecting activities.

Let us now consider the incentives facing income earners. Although convicted tax offenders are required to pay fines for tax evasion, we assume that there is no reward for compliance with tax laws. Hence, medium income earners have no incentive to report high income. Because these individuals always report truthfully, their after tax income is \( Z_m = (1 - t_m)Y_m \).

Consistent with developing country experience, we postulate that high income earners may be able to avoid punishment for tax evasion even when they are audited or inspected. For simplicity, we assume that high types can be ordered according to their “innate” ability of avoiding punishment for tax evasion. For an individual \( b \in [0, p^a_h] \), let \( b \) denote his innate ability
of punishment avoidance. Hence, individual \( b = 0 \) has the lowest ability and individual \( b = P_r \) has the highest ability. Similarly, for an individual \( w \in [0, P_r] \), let \( w \) denote his innate ability of punishment avoidance.

When the TA’s quality rate is \( q \), an audited individual \( w \), who has ability of punishment avoidance \( w \), is punished with probability:

\[
\mu^w(q, w) = q \left( P_r - w \right)
\]

Hence, \( \mu^w(0, w) = \mu^w(q, P_r) = 0 \) and \( \mu^w(q, 0) = q \).

It seems reasonable to assume that, everything else equal, it is harder to punish a high black than a high white. We model this increased hardship with the assumption that the TA is less informed about the black sector than about the white sector. Let \( \kappa \), \( 0 < \kappa < 1 \), denote the degree of useful hard tax information about the black sector available to the TA. When the TA’s quality rate is \( q \), an audited or inspected individual \( b \), who has ability of punishment avoidance \( b \), is punished with probability:

\[
\mu^b(q, b) = \mu(q, b; P_r) \kappa = q \left( P_r - b \right)
\]

Each high white individual faces three possible contingencies: (1) reports \( Y_M \) does not face punishment; (2) reports \( Y_M \) faces punishment; or (3) reports \( Y_H \). Each high black, on the other hand, faces five possible contingencies: (1) reports \( Y_M \) does not face punishment; (2) reports \( Y_M \) faces punishment; (3) reports \( Y_H \); (4) does not file / does not face punishment; (5) does not file / faces punishment. Let \( \sigma \) denote a contingency and \( \Sigma = \{1, 2, 3\} \) denote the set of contingencies for high whites. Define \( Z^w(\sigma) \) as the consumption level of a high white when contingency \( \sigma \in \Sigma \) occurs. Similarly, let \( \Sigma = \{1, 2, 3, 4, 5\} \) denote the set of contingencies for high blacks, and define \( Z^b(\sigma) \) as the consumption level of a high black when contingency \( \sigma \in \Sigma \) occurs. Let \( f \) denote the fine rate for underreporting and \( \phi \) denote the fine rate for not filing a tax return. Then, the consumption levels for high blacks and whites are as follows:

\[
\begin{align*}
Z^w(1) &= Z^w(1) = Y_M - t_M Y_M; \\
Z^w(2) &= Z^w(2) = [1 - (1 + f)t_M]Y_M + f t_M Y_M; \\
Z^w(3) &= Z^w(3) = (1 - t_M)Y_M; \\
Z^w(4) &= Y_M; \\
Z^w(5) &= [1 - (1 + \phi)t_M]Y_M.
\end{align*}
\]

We assume that \( 0 < f < \phi \). To simplify the analysis, let \( \phi = \frac{1 - t_M}{t_M} \). Hence, \( Z^w(5) = 0 \).
Let $U$ represent the utility function for an income earner. For most of what follows, we assume that $U(0) = 0$; $U' > 0$; $U'' < 0$. The expected utility for a high type who reports truthfully is simply $U((1-t_H)Y_H)$. The expected utility for an individual $w$ who reports medium income is:

$$U(Y_m - t_m Y_m) - \left( \frac{A_m}{R_m} \right) q \left( \frac{P^n_m - w}{P^n_m} \right) \left[ U(Y_m - t_m Y_m) - U((1-(1+f)t_H)Y_H + t_m Y_m) \right]$$

Then, individual $w$ reports truthfully if and only if

$$U((1-t_H)Y_H) \geq U(Y_m - t_m Y_m) - \left( \frac{A_m}{R_m} \right) q \left( \frac{P^n_m - w}{P^n_m} \right) \left[ U(Y_m - t_m Y_m) - U((1-(1+f)t_H)Y_H + t_m Y_m) \right]$$

(1)

It is important to observe that condition (1) is violated at $w = P^n_m$. It follows that we are always guaranteed to have $R^n_m < P^n_m$.

The expected utility for an individual $b$ who does not file is

$$\left[ 1 - q K \left( \frac{P^n_b - b}{P^n_b} \right) \right] U(Y_b)$$

The expected utility for an individual $b$ who reports medium income is

$$U(Y_m - t_m Y_m) - \left( \frac{A_m}{R_m} \right) q \left( \frac{P^n_m - b}{P^n_m} \right) \left[ U(Y_m - t_m Y_m) - U((1-(1+f)t_H)Y_H + t_m Y_m) \right]$$

Then, individual $b$ files a tax return if and only if

$$\left[ 1 - q K \left( \frac{I}{N} \right) \right] \left( \frac{P^n_b - b}{P^n_b} \right) \left[ U(Y_m) \right] \leq \max \{ U((1-t_H)Y_H), U(Y_m - t_m Y_m) \} - (q K \left( \frac{A_m}{R_m} \right) \left( \frac{P^n_m - b}{P^n_m} \right) \left[ U(Y_m - t_m Y_m) - U((1-(1+f)t_H)Y_H + t_m Y_m) \right]$$

(2)

Since condition (2) is violated at $b = P^n_b$, we always have $N_H > 0$.

Let $C$ denote the total number of individuals who comply with tax laws. Individuals may comply voluntarily or involuntarily. If $C_v$ denotes the total number of taxpayers who comply voluntarily and $C_o$ denotes the total number of tax law offenders that the court convicts for tax evasion, we have $C = C_v + C_o$. Furthermore, the group of convicted tax evaders consists of those convicted following successful audits, $C_A$, and those convicted following successful inspections, $C_I$. Hence, $C_o = C_A + C_I$. For future reference note also that:

$$C_v = C_v^a + C_v^w = R^n_b + P_m + R^n_m$$

$$C_A = C_A^a + C_A^w.$$
We assume that the government cannot observe either the TA's inputs (efforts) or outputs (audits, inspections and quality level). The government, however, can observe some of the TA's performance outcomes, namely, collections of total tax and fine revenues. This implies that the government can use these indicators to reward the TA's performance. The autonomous TA, therefore, controls its own revenues and allocation of effort levels among its four activities. The TA's objective function is

$$Q = S + p_T T + p_F F + p_D D - \Psi(e_M, e_H, e_U, e_Q, \theta)$$

The quantity $S$ represents the lump-sum transfer received from the government (if positive) or paid to the government (if negative), $T$ denotes the total tax revenue collected, $F$ gives the total fine revenue collected from convicted underreporters, $D$ represents the total fine revenue collected from convicted nonfilers, $p_T$ is the reward rate per dollar of tax revenue collected, $p_F$ is the reward rate per dollar of fine for underreporting collected and $p_D$ is the reward rate per dollar of fine for nonfiling collected. The TA takes the reward rates as given; they are predetermined by the government.

III. The Income Tax Game

We are now ready to describe the “income tax game” played between income earners and the TA. The game is as follows:

The Income Tax Game

Stage 0: The government announces the tax structure – i.e., tax rates, fine rates and performance rewards.

Stage 1: Income earners decide whether or not to file, and those who file decide whether or not to file truthfully.

Stage 2: The TA decides how much effort to exert in each of its auditing and inspecting activities.

In order to focus on the incentives governing the decisions of income earners and the TA, we do not model the government as a strategic player. We assume that the tax structure is determined exogenously. This does not seem to be a strong assumption in what regards tax and fine rates, since in democratic societies they are determined jointly by the executive and legislative branches of government. Performance rewards, on the other hand, may be determined solely by the executive branch. Our assumption here is based on the Brazilian experience with performance bonuses. While the bonus scheme was designed by the Brazilian TA itself, its implementation depended on approval of both the executive and the legislative. A model that correctly depicts the Brazilian case, for example, would need to consider the lobbying efforts exerted by the TA as well as the strategic motivations of both the executive and legislative branches. This is an interesting avenue for future research.
The equilibrium concept used here is subgame perfection. As it is usually done, we start by examining the last stage of the game. First, note that

\[ T \equiv T_m + T_n = t_mR_m + t_nR_n + t_nC_n = t_mR_m + t_nR_n + t_n(C_{\lambda} + C_{r}) \]

\[ F \equiv f_mC_{\lambda}, \quad \Phi = \phi_nC_i = (1 - t_n)C_i. \]

Then, the TA's objective function becomes:

\[ \Omega = S + p_r(t_mR_m + t_nR_n) + t_h\left[ (\rho_r + \rho_f)C_{\lambda} + (\rho_r + \rho_s(1 - t_n))C_i \right] - \Psi(e_m, e_n, e_t, e_o; \theta). \]

Let \( \rho = \rho_r + \rho_f \) and \( \rho_i = \rho_r + \rho_s(1 - t_n) \). Hence:

\[ \Omega = S + p_r(t_mR_m + t_nR_n) + t_h\left[ \rho C_{\lambda} + \rho C_i \right] - \Psi(e_m, e_n, e_t, e_o; \theta). \]

It is important to notice that at the last stage of the game some quantities have already been determined in the previous stages and are thus taken as given by the TA. These quantities are the tax and fine rates, the performance rewards and the numbers of individuals who do not file, who file medium-income reports, \( R_m \), and who file high-income reports, \( R_h \). The TA is mainly interested in finding tax evaders – underreporters and nonfilers – who will not be able to avoid punishment for tax evasion.

It is straightforward to calculate the expected numbers of tax offenders who are convicted for tax evasion. These quantities are given by

\[ C_{\lambda}(e_m, e_n, e_t, e_o) = \left( \frac{k\alpha}{E} \left( \frac{R_m}{R_{\lambda}} \right) \left( \frac{R_n}{P_n} \right) \left( N_m + \frac{R_m^2}{2} \right) \right) \]

\[ C_{\lambda}(e_m, e_n, e_t, e_o) = \left( \frac{e_o}{E} \left( \frac{R_m}{R_{\lambda}} \right) \left( \frac{R_n}{P_n} \right) \left( R_m - R_{\lambda} + \frac{R_m^2}{2} \right) \right) \]

\[ C_i(e_m, e_n, e_t, e_o) = \left( \frac{\alpha}{E} \left( \frac{N_m}{2} \right) \left( \frac{N_n}{P_n} \right) \right) \]

Note that each of the functions above decreases in \( e_n \). This is easy to explain. Auditing high income reporters has no direct deterrent effect on underreporting or nonfiling behavior, but it competes resources away from auditing of medium income reporters and inspecting of nonfilers. Hence, the TA does not find it desirable to exert effort in this activity:

\[ e_n = 0. \]

Let

\[ V(e_m, e_t, e_o) = t_h[p_lC_{\lambda}(e_m, e_n, e_t, e_o) + p_lC_i(e_m, e_t, e_o)] - \Psi(e_m, e_t, e_o; \theta). \]

The TA chooses nonnegative \( \{e_m, e_t, e_o\} \) to maximize (4). To illustrate the circumstances under which both auditing of medium income reporters and inspecting of nonfilers are strategic complements to enforcement quality, it is instructive to solve this maximization problem in two
steps. We shall first keep \( e_0 \) fixed and maximize (4) with respect to \( e_m \) and \( e_i \). This procedure will enable us to define \( e_m \) and \( e_i \) as implicit functions of \( e_0 \) and then characterize the circumstances under which both functions increase in \( e_0 \). In the second step, we place the implicit functions into (4) and then maximize the implied function with respect to \( e_0 \).

Assuming an interior solution, the first order conditions with respect to \( e_m \) and \( e_i \) are as follows:

\[
V_m = \frac{\partial V}{\partial e_m} = t_n \left[ \frac{\rho_n C_A (e_i + e_0)}{E e_m} - \frac{\rho_n C_i}{E} \right] - (1-\theta) e_m = 0, \tag{5a}
\]

\[
V_i = \frac{\partial V}{\partial e_i} = t_n \left[ \frac{\rho_n C_i (e_m + e_0)}{E e_i} - \frac{\rho_n C_A}{E} \right] - (1-\theta) e_i = 0. \tag{5b}
\]

Equations (5) indicate that auditing and inspecting efforts are chosen at levels that equate marginal revenues to marginal costs. Given \( e_0 \), these conditions are both necessary and sufficient for a unique local maximum, since

\[
V_{mm} = \frac{(1-\theta)(2e_m + E)}{E} < 0, \tag{6a}
\]

\[
|H_{mm}| = V_{mm}V_{ii} - V_{mi}^2 = \frac{(1-\theta)[2E(e_i + e_m) + (2e_m + e_0)(2e_i + e_0)]}{E^2} > 0. \tag{6b}
\]

Conditions (6) tell us that the second order leading principal minor of the Hessian matrix of \( V \), \( |H_{mm}| \), is locally negative definite. Hence, we know that equations (5) can be used to implicitly define the functions

\[
e_j = e_j(e_0, \rho_n, \rho_i, t_n, \theta, \alpha_m, \alpha_i, \kappa, R_m, R_i, N_m, P_m, P_i, p_m, p_i, \lambda) \quad j = 1, M.
\]

Differentiating equations (5) with respect to \( e_0 \) yields:

\[
(1-\theta)^2E^2e_0\left(2e_m + E \quad e_m + e_i \quad 2e_i + E\right) \begin{pmatrix}
\frac{\partial e_m}{\partial e_0} \\
\frac{\partial e_i}{\partial e_0}
\end{pmatrix} = \begin{pmatrix}
(1-\theta)(e_m + e_i)E e_m + t_n(\rho_n C_A + \rho_i C_i)e_0 \\
(1-\theta)(e_m + e_i)E e_i + t_n(\rho_n C_A + \rho_i C_i)e_0
\end{pmatrix} \tag{7}
\]

Close inspection of the linear system of equations (7) reveals that at least one of its variables, \( \frac{\partial e_m}{\partial e_0} \) and \( \frac{\partial e_i}{\partial e_0} \), must be positive. In other words, at least one of the enforcement quantities must be strategically complement to enforcement quality. Indeed, it is quite reasonable to expect that both auditing and inspecting are strategic complements to quality, since

\[
\frac{\partial e_m}{\partial e_0} = \frac{(1-\theta)E(e_m + e_i)[\rho_n C_A (E + e_i) - e_m^2] + t_n(\rho_n C_A + \rho_i C_i)[2e_i + e_0]}{(1-\theta)Ee_0(E^2 + (E + e_0)(e_m + e_i) + 4e_m e_i)}, \tag{8a}
\]

\[
\frac{\partial e_i}{\partial e_0} = \frac{(1-\theta)E(e_m + e_i)[\rho_i C_i (E + e_i) - e_i^2] + t_n(\rho_n C_A + \rho_i C_i)[2e_m + e_0]}{(1-\theta)Ee_0(E^2 + (E + e_0)(e_m + e_i) + 4e_m e_i)}. \tag{8b}
\]
Partial derivatives (8) are both positive if \( e_M = e_i \). They are also positive with unequal auditing and inspecting efforts provided these quantities are not strongly dissimilar. It is possible that one effort level is so large relative to the other that one of the partial derivatives turns out to be negative. For example, (8a) could be negative if \( e_i \) is very large relative to \( e_M \). The other partial derivative, however, would necessarily be positive. To minimize ambiguities in what follows, we shall assume that, for any \( e_0 \geq 0 \), the optimal auditing and inspecting efforts chosen by the TA satisfy the following inequalities:

\[
\begin{align*}
  e_M(E + e_i) &> e_i^1, \\
  e_i(E + e_m) &> e_m^2.
\end{align*}
\]

Although inequalities (9) are apparently quite reasonable, in the sense that they hold as long as auditing and inspecting efforts are not too strongly dissimilar, it should be clear that they are only sufficient, not necessary, conditions for the partial derivatives (8) to be positive. Inequalities (9) are quite useful because they enable us to sign some partial derivatives in our comparative statics analysis below.

To examine the TA's choice of quality effort, let us define \( \Gamma(e_0) = V(e_m(e_0), e_i(e_0), e_0) \). Maximizing \( \Gamma(e_0) \) with respect to \( e_0 \) and assuming an interior solution we obtain:

\[
\Gamma_0 = \frac{d\Gamma}{de_0} = \frac{t_h(p_c e_m + p_i e_i)(e_m + e_i) - (1 - \delta) e_0}{E e_0} = 0.
\]

Equation (10) tells us that the TA exerts quality effort at a level that equates the marginal revenue and marginal cost from such an activity. It is important to alert the reader that, due to the strategic complementarities, \( \Gamma \) is not necessarily strictly concave in the neighborhood of a maximum. Fortunately, the (omitted) condition required for \( \Gamma \) to be locally strictly concave does not appear to be very restrictive. Assuming that this condition is satisfied, we can use (10) to implicitly define the following function:

\[
e_0 = e_0(p_c, p_i, t_h, \delta, \alpha_m, \alpha_i, \kappa, R_m^s, R_m^w, N_m, P_t, P_m, P^s, P^w).
\]

At stage 1 of the game, income earners anticipate how the TA will react to their decisions and behave accordingly. Given the results above, conditions (1) and (2) may be rewritten as follows:

\[
U(t - t_h)Y_m \geq U(Y_m - t_m Y_m) - \left( \frac{\alpha_m e_m(e_0(e_0))}{R_m}\right) \left( \frac{e_0(e_0)}{E(e_0)} \right) \left( \frac{P_m}{P^w} - \frac{w}{w_r} \right) \left[ U(Y_m - t_m Y_m) - U([1 - (1 + t_h)]Y_m + t_m Y_m) \right]
\]
\[
\left[ 1 - \left( \frac{\alpha_e c_0(\epsilon_0)\epsilon\kappa}{N}\right) \left( x_{c_0(\epsilon)}(\epsilon) \right) \left( \frac{P_m - b}{E(\epsilon)} \right) \left( \frac{P_m^{\ast} - b}{P_m^{\ast}} \right) \right] \left[ U(Y_m) - \max\{U((1 - t_m)Y_m), U(Y_m - t_m Y_m) \}ight] - \left( \frac{\alpha_e c_m(\epsilon_0)\epsilon}{R_m^{\ast}} \right) \left( x_{c_m(\epsilon)}(\epsilon) \right) \left( \frac{P_m^{\ast} - P_m}{E(\epsilon)} \right) \left( \frac{P_m^{\ast}}{P_m^{\ast}} \right) \left[ U(Y_m - t_m Y_m) - U((1 - (1 + f)t_m)Y_m + \alpha_m Y_m) \right],
\] 

(12)

where \( E(\epsilon) = c_m(\epsilon_0)\epsilon + c_e(\epsilon_0)\epsilon + c_0(\epsilon) . \)

We now investigate the properties of an equilibrium for an economy characterized by a low level of development. We define such an economy as one where we observe a low enough value for the parameter \( \kappa \) so that condition (12) is violated for all \( b \in [0, p_m^{\ast}] \). This, in turn, implies that \( N_h = p_m^{\ast} \) and \( R_m^{\ast} = R_m \). In words, all high blacks are nonfilers and thus those who report medium income are all whites, either medium or high whites. Furthermore, we assume that condition (11) is satisfied binding for some type \( w' \in [0, p_m' \] when \( N_h = p_m^{\ast} \) and \( R_m^{\ast} = R_m \). Hence, all high whites whose types are in the interval \( [0, w'] \) report truthfully and all high whites whose types are in the interval \( [w', p_m' \] underreport their incomes. That is to say, for \( w = R_m^{\ast} = p_m' = (R_m^{\ast} - p_m) \) and

\[
- \left( \frac{\alpha_e c_0(\epsilon_0)\epsilon}{R_m^{\ast}} \right) \left( x_{c_0(\epsilon)}(\epsilon) \right) \left( \frac{P_m^{\ast} - P_m}{E(\epsilon)} \right) \left( \frac{P_m^{\ast}}{P_m^{\ast}} \right) \left[ U(Y_m - t_m Y_m) - U((1 - (1 + f)t_m)Y_m + \alpha_m Y_m) \right],
\] 

(13)

It is important to note that the quantity \( R_m^{\ast} - p_m \) gives us the number of high whites who underreport income; i.e., since \( p_m \) is fixed, changes in \( R_m^{\ast} \) represent changes in the number of underreporters. Equation (13) defines \( R_m^{\ast} \) as an implicit function of the relevant exogenous variables:

\[
R_m^{\ast} = R_m^{\ast}(t_m, t_H, Y_m, Y_H, \epsilon, \rho_X, \rho_I, \beta, \alpha_m, \alpha_e, \kappa, p_m, p_L, p_m')
\]

It is useful to summarize the results that we have obtained so far. The subgame perfect equilibrium for the income tax game played by income earners and the TA is characterized by equations (3), (5a), (5b), (10), (13), \( N_h = p_m^{\ast} \) and \( R_m^{\ast} = R_m \). Equations (5a), (5b) and (10) inform us that the efforts exerted by the TA in auditing medium-income reports, inspecting nonfilers and improving quality are set at levels that equate the TA's marginal revenues to its marginal costs. Equation (3) tells us that the TA finds it unprofitable to exert any effort at all in audits of high-income reports. Equations (13), \( N_h = p_m^{\ast} \) and \( R_m^{\ast} = R_m \) describe the equilibrium for income earners in an economy characterized by low development, an economy where the degree of information about the black sector is so low relative to the degree of information available about the white sector that no high black finds it advantageous to file a tax return. Equation (13), however, demonstrates that the TA's efforts are not totally in vain, since they deter some high whites from underreporting their incomes.

The strategic complementarity there necessarily exists between at least one of the effective enforcement quantities and enforcement quality is, perhaps, the most important message of our analysis up to this point. Knowing that improvements in enforcement quality lead
to improvements in audits and inspections, the TA does not have any incentive to sacrifice enforcement quality in order to save resources that can be allocated in the other activities. Conditions (8a) and (8b) demonstrate that at least one of the effective enforcement quantities, audits of medium-income reports and inspections of nonfilers, must be strategically complement to enforcement quality. Conditions (9a) and (9b), which rule out strongly dissimilar effort levels, provide sufficient conditions for both effective enforcement quantities to be strategic complements to enforcement quality.

IV. Comparative Statics

It is important to know which refutable hypotheses, if any, are implied by our model. As we shall discuss below, our comparative statics analysis is plagued with ambiguities. Whenever possible, we will attempt to shed some light on the sources of ambiguities. In a few cases, the relationships become unambiguous by further placing some apparently reasonable restrictions.

To better understand the sources of ambiguities, we provide a list of the relevant partial derivatives for each stage of the income tax game. To save space, we do not present the partial derivatives per se, but only indicate if they are positive, negative or ambiguous. Consistent with the way we solved the game, we start by listing the relevant partial derivatives for the last stage of the game. Given \( e_0 \), we obtain the following partial derivatives for the enforcement quantity efforts:

\[
\begin{align*}
\frac{\partial e_{M}}{\partial \rho_{\lambda}} & > 0; & \frac{\partial e_{1}}{\partial \rho_{\lambda}} & < 0; & (14a) \\
\frac{\partial e_{M}}{\partial \rho_{1}} & < 0; & \frac{\partial e_{1}}{\partial \rho_{1}} & > 0; & (14b) \\
\frac{\partial e_{M}}{\partial \theta} & > 0; & \frac{\partial e_{1}}{\partial \theta} & > 0; & (14c) \\
\frac{\partial e_{M}}{\partial \alpha_{M}} & > 0; & \frac{\partial e_{1}}{\partial \alpha_{M}} & < 0; & (14d) \\
\frac{\partial e_{M}}{\partial \alpha_{1}} & < 0; & \frac{\partial e_{1}}{\partial \alpha_{1}} & > 0; & (14e) \\
\frac{\partial e_{M}}{\partial \kappa} & < 0; & \frac{\partial e_{1}}{\partial \kappa} & > 0; & (14g) \\
\frac{\partial e_{M}}{\partial R_{M}} & > 0; & \frac{\partial e_{1}}{\partial R_{M}} & < 0; & (14h) \\
\frac{\partial e_{M}}{\partial P_{c}} & > 0; & \frac{\partial e_{1}}{\partial P_{c}} & < 0; & (14i) \\
\frac{\partial e_{M}}{\partial P_{m}} & > 0; & \frac{\partial e_{1}}{\partial P_{m}} & < 0; & (14j) \\
\frac{\partial e_{M}}{\partial P_{n}} & > 0; & \frac{\partial e_{1}}{\partial P_{n}} & < 0. & (14k)
\end{align*}
\]
\[ \frac{\partial e_m}{\partial p_A^w} < 0; \quad \frac{\partial e_I}{\partial p_A^w} > 0. \] (14I)

Partial derivatives (14a) tell us that auditing effort increases and inspecting effort decreases as the "auditing performance reward rate," \( p_A \), increases. Auditing and inspecting activities are strategic substitutes, even though each activity is strategically complement to enforcement quality. The rationale for partial derivatives (14b) is identical to the rationale for partial derivatives (14a). The signs of partial derivatives (14c) and (14d) follow from our assumption that enforcement quantity efforts are not strongly dissimilar, conditions (9). Audits of medium reports and inspections of nonfilers become more lucrative as the highest tax rate increases. Provided the TA does not strongly favor one of these activities over the other, effort levels increase in both activities. By a similar reasoning, the profitability of both activities rises when the overall cost of production lowers as a result of an improvement in the degree of computerization.

An improvement in auditing productivity leads to an increase in auditing effort and a decrease in inspecting effort, as partial derivatives (14e) indicate. The rationale for partial derivatives (14f) is identical. As for the effects on enforcement quantity efforts of an improvement in the TA's degree of information about black sector activities, we find that inspecting effort definitely rises because inspections of nonfilers becomes more lucrative. In principle, it is not clear whether such an improvement would increase or decrease the profitability of audits of medium reports. If some high blacks were to file returns, auditing of medium reports would become more profitable, holding everything else constant. But, the opportunity cost of increasing audits may be substantial. Since auditing and inspecting are strategic substitutes, an increase in audits would necessarily compete resources away from inspections. Hence, provided some high blacks filed returns, the information improvement would produce positive and negative effects on auditing of medium reports, and the implied relationship between these quantities would be ambiguous. In the equilibrium examined in this paper, however, no high black files a return. This implies that the TA finds it desirable to reduce auditing effort as the degree of information about the black sector improves.

To understand partial derivatives (14h) - (14l), it is instructive to remember how the expected numbers of people convicted for underreporting and nonfiling are calculated, functions \( C_A \) and \( C_I \). Since, in equilibrium, \( R_A = 0 \), we have \( C_A = 0 \) and \( C_A = C_I \). Now, note that an increase in the number of medium reports by high whites leads to an increase in \( C_A \) and to no change in \( C_I \). Therefore, auditing of medium reports becomes more profitable relative to inspecting of nonfilers as the number of underreporters increases, effects that partial derivatives (14h) illustrate. Since the logic for results (14i) - (14l) is similar, we will not comment on them.

The results concerning the comparative statics for enforcement quality are murky. Because both enforcement quantity efforts depend positively on enforcement quality, the effects on enforcement quality effort of changes in the parameters are both direct and indirect. As the indirect effects influence the overall effects as a sum, the overall effects on enforcement quality effort cannot be unambiguously determined unless both indirect effects have the same sign. For example, the effect on enforcement quality effort of an increase in the auditing performance reward rate is ambiguous because such an increase leads to a simultaneous expansion in auditing
effort and contraction in inspecting effort, and the overall effect on enforcement quality effort depends, among other things, on the sum of these two effects. Therefore, there are only two partial derivatives for the enforcement quality effort function that can be unambiguously signed:

\[
\begin{align*}
\frac{\partial e_0}{\partial t} &> 0 & \text{and} & \frac{\partial e_0}{\partial \theta} > 0.
\end{align*}
\] (15)

The first partial derivative in (15) tells us that the TA increases quality effort as the highest tax rate rises because such a change in highest tax rate raises the profitability of enforcement quality. This result is implied by the positive impact on the profitability of both enforcement quantities following an increase in the highest tax rate. By a similar token, an improvement in the degree of computerization leads to a decrease in the overall cost of production. This motivates the TA to exert more effort on both enforcement quantity activities and, consequently, to increase enforcement quality effort.

Let us now turn to the relevant comparative statics for stage 1 of the game. We shall examine the effects on the variable \( R^*_w \) of changes in the set of parameters \( \{ \theta, t, t_m, Y, Y_m \} \). As we observed above, changes in the other exogenous variables lead to ambiguous results. Unfortunately, none of the signs of the partial derivatives for the \( R^*_w \) function can be determined unambiguously. The sources of ambiguities in this case are basically two: (i) the concavity of \( U \) (i.e., risk aversion); and (ii) the enforcement quality rate, \( q = e_0/E \).

Risk aversion is a source of ambiguity for all partial derivatives except for the partial derivative with respect to \( 2 \). The only source of ambiguity for this partial derivative concerns the effect on the enforcement quality rate of changes in \( 2 \):

\[
\frac{\partial q}{\partial \theta} = \frac{1}{E} \left( \frac{\partial e_0}{\partial \theta} - \frac{1}{E} \frac{\partial E}{\partial \theta} \right) \] (16)

Since \( \frac{\partial e_0}{\partial \theta} > 0 \) and \( \frac{\partial E}{\partial \theta} > 0 \), the sign of partial derivative (16) cannot be determined unambiguously. This is problematic because equation (13) clearly demonstrates that the variable of interest, \( R^*_w \), depends on the enforcement quality rate.

If we assume, as it seems reasonable, that the enforcement quality rate increases with an improvement in the degree of computerization, \( \frac{\partial q}{\partial \theta} > 0 \), we are able to conclusively state that underreporting falls as the TA becomes more computerized, \( \frac{\partial R^*_w}{\partial \theta} < 0 \). The TA's optimal responses to an improvement in the degree of computerization are effective in increasing the expected cost of underreporting. Consequently, fewer high whites attempt to evade taxes.

If we assume risk neutrality instead of risk aversion, we obtain conclusive results for three of the remaining four partial derivatives. With risk neutral income earners, equation (13) simplifies to:
Unlike before, the implicit function defined by (17) is independent of the parameters \( \{t_M, Y_M, Y_u\} \).

These parameters do not influence the TA's efforts. It follows that a change in any of these parameters does not affect the expected cost of underreporting. Hence:

\[
\frac{\partial R^n_w}{\partial t_M} = \frac{\partial R^n_w}{\partial Y_M} = \frac{\partial R^n_w}{\partial Y_u} = 0.
\]  

(18)

Similar to the case for the 2 parameter, we can only determine the sign of the partial derivative with respect to \( t^n_1 \), \( E^n_M \), if we further assume that:

\[
\frac{\partial q}{\partial t^n_1} = 1 \left( \frac{\partial c_0}{\partial t^n_1} - \frac{1}{E} \frac{\partial E}{\partial t^n_1} \right) > 0.
\]  

(19)

If the enforcement quality rate is positively related to the highest tax rate, underreporting falls as the highest tax rate rises. The TA's optimal responses to an increase in the highest tax rate raise the expected cost of underreporting which, in turn, effectively deters some high whites from underreporting their incomes.

V. Conclusion

The simple model studied in this paper illustrates some of the positive incentive effects of tax performance outcome indicators. By rewarding total collections, the government induces the autonomous tax administration to exert effort in improving the quality of tax enforcement. Improvements in quality effort, in turn, may increase the efficiency of both auditing of underreporters and inspecting of nonfilers. Although audits of underreporters and inspections of nonfilers are substitutes for each other, at least one of these enforcement quantities is necessarily positively related to enforcement quality.

Holding enforcement quality effort constant and assuming that the enforcement quantity efforts are not strongly dissimilar, we were able to sign most of the relevant partial derivatives in our comparative statics analysis for the quantity effort variables. These partial derivatives had the expected signs. The two most important results turned out to be the impacts on the quantity effort variables of changes in the high tax rate and in the degree of computerization. We showed that quantity efforts rise with increases in the high tax rate and in the degree of computerization. These results were quite significant because they enabled us to conclusively ascertain that enforcement quality also rises with increases in such parameters.

Due to ambiguities introduced by risk aversion and the enforcement quality rate, we were unable to obtain conclusive results in our comparative statics analysis concerning tax evasion incentives. By restricting our attention to situations where the enforcement quality rate rises with
increases in the degree of computerization, we were able to state that tax evasion falls with improvements in computerization. Tax evasion also falls with increases in the highest tax rate if income earners are risk neutral and the enforcement quality rate rises with increases in the highest tax rate. Both of these effects on tax evasion incentives are due to the actions undertaken by the TA in response to increases in the degree of computerization and in the highest tax rate. Because each of such changes in the parameters motivates the TA to work harder in all enforcement activities, the expected cost of underreporting increases. Hence, fewer high whites attempt to evade taxes.
References


Chapter 8
Leadership or Institutions?
Reputation, Retaliation and Opportunistic Taxation*

Margaret S. McMillan

I. Motivation/Overview

What kinds of leadership are associated with successful tax regimes? What kinds of institutions restrain leaders from predatory taxation? Once a government has established a reputation as predatory, are there institutions that can help to alter investors' perception of the government's reputation?

McMillan (2000) establishes the level of output taxation as the equilibrium outcome of an infinitely-repeated game between government officials and the private sector. When private investment is irreversible, predatory governments sometimes (but not always) succumb to the temptation to tax away sunk investments. McMillan and Masters (2000) extend the model to show that the same set of institutions that lead to opportunistic taxation (for example, high sunk costs and a high discount factor) also lead to underinvestment in public goods (research and development spending in our model) and ultimately low growth. The reason for the underinvestment in public goods is once again a short time horizon (since the rewards from investing in R&D are only realized in future periods) and little or no consequences of not doing so. The present paper summarizes these results and explores their institutional implications for tax design and reform taking into account both the limitations of the theoretical model and data constraints.

Our key finding is that strong leadership (or reputation) unchecked by institutions for retaliation enables opportunistic taxation. In prescribing solutions, we pay particular attention to the fact that countries differ in their institutional endowments, and so will necessarily differ in their abilities to adopt and implement the proposed solutions. This work draws on the time-consistency concept of Kydland and Prescott (1977), and the applications to intertemporal price policy of Besley (1997), McLaren (1996) and Gilbert and Newbery (1994).

* This is a revised version of a paper presented at the Columbia University-World Bank Conference on Institutional Elements of Tax Design and Reform, February 18-19, 2000, at Columbia University. I would like to thank Ken Leonard, John McLaren, William Masters and seminar participants for helpful comments.
II. The Model

In our approach, farmers choose between remaining at subsistence and producing for market, where they can earn positive profits but are exposed to taxation. The government controls the marketplace, and sets the tax by paying farmers less than the price received from consumers which we assume to be an exogenous world price.

We assume that policymakers have an infinite time horizon, and seek to maximize the present discounted value of some social welfare function which is a weighted sum of tax revenue and producer surplus. In an extreme case where policymakers place no value on producer surplus, the tax that maximizes their objective function is the revenue maximizing tax. But such a tax would not be sustained in equilibrium if private investment is irreversible and the government cannot fix future tax rates.

Without credible commitment, a government that discounts producer surplus could be tempted to announce low taxes to induce investment, then raise taxes after sunk costs are incurred to expropriate the resulting economic rent. Thus, observed investment levels and tax rates are part of a repeated game between policymakers and farmers, whose equilibrium depends on the incentives available in deviation (or defection) from the optimal (or low-tax) policy.

To determine whether the low-tax policy is sustainable as an equilibrium, we must be more specific about the consequences of deviating from the low-tax policy. We assume that there are many small farmers and that these farmers would prefer the low-tax policy to any other opportunity and so have no incentive to defect unless the government does not honor its commitment. The government may be tempted to deviate from the low-tax policy to gain revenue. If the government does deviate and expropriates farmers’ sunk costs, each farmer individually learns from this experience and makes no new investments for a number of periods, $k$. Thus, although the farmers are not formally organized, each one acting in their own self-interest generates a coherent retaliation strategy for farmers in the aggregate.

Conditions for sustaining a low-tax policy are derived by comparing the payoffs to the government of sticking with the low-tax policy to those associated with defection\(^1\). This comparison yields the following expression:

\[
(1)\quad (1 - \alpha)STC \leq \delta(T,k)\pi^*(P_f, P_w, rd)
\]

The left-hand side of inequality (1) represents the government’s short-run gains available from imposing confiscatory taxes, in defection away from the low-tax equilibrium: it is the ratio of sunk costs to total costs (STC), weighted by government’s relative valuation of farm income as opposed to tax revenue ($\alpha$). The right-hand side is the present discounted costs of doing so, or the long-run tax revenue foregone by deviating from the low-tax path. We call this expected profitability ($\pi^*$) because it depends on expected future world prices and the future productivity.

\(^1\) More detail on the derivation of this inequality can be found in McMillan and Masters (2000).
of spending on research and development \( (P', P^W, rd) \). The length of time over which discounting \((\delta)\) occurs depends on the period of time over which production continues without further investment \((T)\), and also the period over which farmers revert to subsistence after suffering confiscatory taxes \((k)\).

The equilibrium condition above provides the testable hypotheses of the model. The factor highlighted in our model that is generally omitted from other analyses is \( STC \), the observed ratio of sunk to total costs. The higher is this \( STC \) ratio, the greater is the government's incentive to undertake predatory taxation after investment occurs. The \( STC \) ratio is largely a physical characteristic of production technology, and is relatively high for perennials and other crops requiring a heavy up-front investment in irrigation or field preparation. Tropical countries in general, and African ones in particular, tend to have a comparative advantage in these products, particularly tree crops like coffee and cocoa, perhaps because their climate offers less of a concentrated summer growing season with less available moisture and generally poorer soils than temperate regions. By favoring crops requiring heavy pre-harvest investment, the physical environment itself can be said to make farmers relatively more vulnerable to predatory taxation, thus inhibiting growth unless governments can commit to low-tax policies.

Two other variables, \( a \) and \( \delta(T,k) \) formalize the role of political conditions that are often discussed in previous studies. \( a \) reflects the degree to which the government is representative of farmers as opposed to those who benefit from tax revenue. For example, Bates (1981) argues that \( a \) is fairly low because farmers are not organized enough to pose a threat to the government in power. \( \delta(T,k) \) reflects the degree to which the government is impatient and discounts future tax revenues. \( T \) represents the length of time over which a crop continues to produce before it must be replanted. For annual crops, \( T \) is equal to one and for perennial crops, \( T \) will vary depending upon when the crop was actually planted. The parameter \( k \) represents the length of time over which farmers refuse to plant once they've been cheated out of sunk costs, or the strength of the farmers' retaliation.\(^2\)

The remaining two variables, expected future world prices and the productivity of R&D, are subject to debate. However, recent projections of Africa's terms of trade (Hertel and others 1998) suggest continued high demand for African farm products, and studies of the productivity of African agricultural research suggest that it is at least as productive as research elsewhere (Masters and others 1998).

The political-economy model described above provides testable predictions about which countries will adopt what policies. In the cross-country empirical implementation, we first ask what determines policy, looking particularly for an effect of cross-country differences in the \( STC \) ratio and cross-country differences in the discount factor.

### III. Empirical Application, Data and Results

The model predicts that a low-tax, high-growth equilibrium will be harder to sustain: (a) the larger the share of sunk costs in total costs, (b) the smaller are expected future profits from a particular investment, and (c) the smaller the government's discount factor. In estimation we will

\(^2\) In this simple model, \( k \) only represents years during which farmers refuse to plant. However, in a model in which the discount factor is endogenized, one could imagine that \( T \) might be a function of formal or informal institutions available for retaliation.
treat each government’s weighting of producer surplus as an unobservable variable taking on country-specific values.

A direct test of the model would require us to classify countries in terms of whether their policies are consistent with maximizing any government objective function, and then regress that regime classification on its determinants. McMillan (2000) provides details of this approach. Here we use a more general version of the model, regressing the magnitude of taxation and R&D investment directly on their determinants and then regressing growth on taxation and R&D. This approach yields easy-to-interpret estimates of elasticities, and is directly comparable to others’ work in the empirical-growth literature.

Of the three variables in equation (1), only $STC$ is directly observable. Though observable, its measurement is non-trivial since data on costs of production in Sub-Saharan Africa are not readily available. These data were collected by me and are published in McMillan (2000). Expected future revenue depends on expected future commodity prices, which depend on the way in which expectations are formed about future world prices and on costs of production. Following Deaton and Miller (1995), I estimate expected future profitability using a twenty-year average of the ratio of world prices to costs of production. Finally, the discount factor, $\delta (T, k)$, is estimated as one over one minus the mean duration in power of past leaders using data from Bienen and Van de Walle (1991). In section V of the paper, we discuss the limitations of this measure of the discount factor.

Our dependent variables as well as some of the explanatory variables in the growth regressions are obtained from secondary sources. Initial income is GDP per capita in purchasing power parity dollars in 1965, from the Penn World Tables version 5.6. Growth is measured as the average annual change in the natural logarithm of GDP per capita between 1965 and 1990. Agricultural R&D is measured in real per-capita terms, and is derived from the work of Pardey et al. (1998). R&D expenditures are available on an annual basis for a total of 19 countries over a period of 30 years, 1961-1991. Agricultural taxation is measured using the nominal protection coefficient (NPC), the farmgate to border price ratio. Thus, an increase in the NPC reflects a reduction in taxation as it represents a reduction in the divergence between what farmers could get if they sold their product directly to world markets and what they actually get due to government intervention. These data are derived from the work of Jaeger (1992), extended by McMillan (2000), and are available for a total of 56 crops and 32 countries for various years. In the growth regressions, the crop-by-crop tax rates are aggregated up to national levels using production weights (McMillan and Masters 2000).

Regression results shown in the table show consistent support for the model, with strong statistical significance for the sunk-to-total-cost ratio, expected profitability, and government’s discount rate in determining agricultural tax rates. The link between underinvestment in public goods and GDP growth is straightforward in our model since the public good is R&D. A failure to invest in R&D reduces total factor productivity and hence leads to little or no growth. Results of this analysis support the theory and are shown in the last two rows of Table 8-1. These results are robust to a wide variety of specification tests, including the main alternative regressors used in the cross-country growth literature, as discussed in McMillan and Masters (2000).
IV. Implications for Institutional Elements of Tax Design

Our key finding is the following: a successful tax system requires both a long-lived regime and this same regime must be vulnerable to swift and effective retaliation. A government without a long time horizon (which is likely to lose power) faces an overwhelming temptation to take the money and run. However, even a government (likely to remain in power) with a long time horizon may not do the right thing if there is no threat of economic retaliation.

The question then is what kinds of institutions will enable a strong leader and at the same time, make that leader vulnerable to retaliation. We discuss this question here in terms of the variables in the model, focusing primarily on the government’s discount factor, $\delta$, and its arguments. The parameter $\alpha$, or weight on producer surplus in the government’s objective function, has a well-established influence—it is clear that taxes will be higher on groups whose political influence is low, but that doesn’t get us very far since those with a low political influence have, by definition, little ability to reform institutions in their favor.

Role of the Government’s Discount Factor, $\delta$

An example of the importance of a long time-horizon (or relatively high probability of remaining in power) is the contrast between the Ivory Coast and Ghana. Our empirical proxy tells us that a leader in Ghana in 1989 would have had a probability of remaining in power equal to .68 while Houphouet-Boigny in the Ivory Coast had a probability of remaining in power approximately equal to one! Thus, ceteris paribus, we would have expected the low-tax equilibrium to be much more likely in the Ivory Coast. In fact, the average annual tax on cocoa in the Ivory Coast over the period 1970-89 was 39 percent while the average annual tax on cocoa farmers in Ghana over this same period was 63 percent.

Role of Investor’s Ability to Retaliate, $k$

Recall that our argument is that a high probability of remaining in power is in and of itself insufficient to guarantee a low-tax regime. This point is best illustrated by comparing the cases of the Ivory Coast and Tanzania: countries with similar production systems and similarly long-lived leaders, but different ideologies associated with different responses to social conflict. In the Ivory Coast the government sought to maintain farmers’ voluntary participation in the market economy while in Tanzania, the government sought to compel farmers into villagization and other socialist experiments. The result was relatively better performance in the Ivory Coast than in Tanzania. Thus, both countries exhibited "strong leadership" or long-lived governments but one had a much more equitable tax system. The comparison of Ivory Coast and Tanzania raises the following question: what else must be in place to insure that a forward looking government does not tax opportunistically?

Based on the simple model laid out in section II, we argue that there must be institutions in place (formal or informal) that provide investors with a serious and credible threat of

---

3 In 1989, Houphouet-Boigny had been in power for 30 years. Thus, the mean duration in power in the Ivory Coast was 30 years and the associated hazard rate in 1989 was .03. The associated probability of remaining in power in 1989 was .97 or approximately 1.
The importance of the possibility of retaliation on leaders actions is described by Jansen in Zambia, who wrote that "taxing the farmer is the most feasible solution from the standpoint of political expedience. Quite simply, the cabinet and members of the central committee of the party know that reducing producer prices—or simply not increasing them to offset a high inflation rate—will not result in riots among the dispersed and relatively less powerful rural constituencies."

Where farmers are politically more active, taxation is less likely: in Nigeria farmers revolted when taxes were too high and successfully got government to lower taxes. A further informal economic institution not described by the model is that of smuggling. Smuggling tends to have a negative connotation yet, paradoxically, in our framework smuggling may discipline the government into sticking with its commitment to a low-tax policy.

The importance of rural insurrection and smuggling points to the importance of retaliation but these are certainly not examples of well-established formal institutions that a country would hope to develop. They are violent and costly means of retaliation that don't always work. For example, in Tanzania, farmers revolted against heavy taxation by burning down coffee plants. These revolts were quieted by the military, that was ordered to imprison these farmers. Other, more effective forms of retaliation are clearly needed.

**Role of Expected Future Profitability, $\pi^e$**

Expected future profitability (for example, expected future tax revenue) depends on the costs of production relative to expected future world prices. Government has two channels through which it can influence expected future profitability. First, government can invest in (or get others to invest in) research and development that increases profitability. A government that is patient or a legal system that protects patents will both tend to increase spending on R&D thus increasing expected future profitability and lowering the relative payoff from deviating to a high tax policy. In addition to these two channels, government can form cartels with other producing nations both to increase and stabilize world prices, but such cartels are typically short-lived—the 1999 collapse of the world rubber cartel (INRO) being a typical example.

**V. Institutional Solutions**

One important implication of the preceding analysis is that the formal economic institution used by government to implement taxation (for example, marketing boards or Caisses de Stabilisation in francophone Africa) is not in and of itself the reason for opportunistic taxation. Perhaps if the model has one important message about economic institutions it is that abolishing marketing boards or privatization is probably not going to solve the problem of predatory taxation. If marketing boards were the only problem, then we would see no variation in tax rates among countries that have marketing boards and that produce similar crops. Yet, we do. Consider again the cases of Ghana and the Ivory Coast. In both countries the economic

---

4 In our model, $k$ represents the severity of the retaliation by farmers or the length of time over which they refuse to plant once cheated out of sunk costs. A richer model might expand the role of farmers providing them with alternative means of retaliation. Note also, that $k$ is not accounted for in the empirical work due to the difficulty of measuring $k$.

5 *West Africa* 1984.
organization of export crop marketing was similar yet Ghana almost taxed its cocoa sector into the ground while cocoa in the Ivory Coast flourished and the government even set aside resources for cocoa research. According to our model, we should have expected this because the government in Ghana has a very high discount factor relative to the government in the Ivory Coast.

So, the question is: what kinds of institutions simultaneously provide leaders with a long time horizon and the threat of retaliation? One example of the type of institution that could provide leaders with a long time horizon and at the same time the threat of retaliation is the political party. The political party, if it works well, disciplines its members by providing or withholding support across generational lines. Since the power of older members depends on their ability to gain and keep the loyalty of younger members, the introduction of the political party helps align the incentives facing individual politicians with the interests of an infinitely-lived dynasty. This, combined with electoral competition among parties, might help discourage the leaders of any one party from emptying the coffers and running off to Switzerland, or even instituting tax regimes that deplete the tax base gradually.

Of course, opposition parties are routinely squashed in many of the countries we are talking about. It is notable that the only two countries in Sub-Saharan Africa that have long-lived opposition parties are Botswana and Mauritius, the two "stars" of Africa. Establishing a strong opposition party is not something that happens over night and many poor countries have weak institutional endowments so the question is what steps can be taken in the interim to balance leadership with institutions for retaliation? Some measures include strengthening the freedom of the press, establishing an independent judiciary and establishing the rule of law. Or, in the words of Hall and Jones, governments must strengthen their anti-diversion policies—not only to encourage private-sector investment directly, but also to restrain government from reversing course in the future.

What else can countries with weak institutional endowments do? One way out of the particular problem described in section II has to do with technology. Citizens are vulnerable to predatory taxation because they have incurred sunk costs. Thus, technical change that raises the importance of variable inputs in production also raises producers' ability to retaliate against predatory taxation. When farmers can make larger and more frequent adjustments in their input use, supply becomes more elastic over time, reducing the revenue available from taxation. A predatory government that attempts heavy taxation will see production fall quickly, leading to

---

6 Technically, Ghana has a marketing board, CMB, while the Ivory Coast has a Caisse de Stabilisation (CSSPPA). The marketing board controls all stages of marketing, from purchases at the farmgate through exports or domestic sales. The Caisse is fundamentally the same as the marketing board in that both own the crop and both play a major role in fixing producer prices and the physical handling of the crop. Unlike the marketing board in Ghana, the Caisse in the Ivory Coast allows private agents to export the crops. (Varangis, Akiyama and Thigpen, 1990)

7 Interestingly, in addition to leadership turnover, Ghana is one of the countries in Sub-Saharan Africa that has had some turnover in the political party of the chief executive (DPI, 2000).

8 A lengthy discussion of political parties is beyond the scope of this paper. Nonetheless, a fair question is, if parties are so useful then why do countries like Italy and Turkey with their multiplicity of parties seem to have so much political instability? According to Dahl (1971), "There are costs of 'toleration' associated with increasing the number of political parties and thus one should seek to avoid a great multiplication of parties."

9 In Botswana, the largest opposition party, Botswana National Front, has been in existence for 17 years and in Mauritius, the largest opposition party MMM, has also been in existence for 17 years. The majority of countries in Africa have only one party, the leader's, and no opposition party. Author's calculations based on the Database on Political Institutions (2000).
less taxation in equilibrium. Some increase in the importance of variable costs happens gradually with intensification, but the development of international R&D for new seeds and better fertilization plays a key role here.

An alternative solution proposed by Levy and Spiller (1993) for the regulation of the telecommunications industry is the substitution of international institutions for missing national institutions. The examples they provide are those of the telecommunications industries in Jamaica and the Philippines. They write:

Jamaica and the Phillipines in the 1950s came closest to using this mechanism. Jamaica's judicial system continues to recognize the Privy Council in London as the final arbiter, conferring continued credibility on its own regulatory system. The Phillipines was granted independence from the United States in 1946, but for the next fifteen years the continuity of its preindependence institutions, the strong influence of the United States, and specific agreements protecting U.S. investors provided a predictable and safe environment that facilitated investment by both Philippine and U.S. investors. Although the specific agreements with the United States have lapsed, the Philippine Long Distance Telephone Company remains subject to some extranational controls as a company with shares traded on the New York Stock Exchange.

In addition, Levy and Spiller suggest that the potential exists to go further in using international institutions as substitutes for domestic weaknesses in commitment capability. One innovation they cite is World Bank guarantees against noncommercial risk, including the risk of administrative expropriation.

Customs unions or regional trade agreements may also limit the scope for discretion. For example, Rodrik (1998), in an attempt to make sense out of Botswana's relative economic success suggests the following:

Along with Lesotho, Swaziland and South Africa, Botswana has long been a member of the South Africa Customs Union (SACU). This means that Botswana has no independent trade policy; goods circulate freely between it and South Africa. The government gets a share of customs revenue collected by South Africa, which amounts to around 20 percent of the value of Botswana's imports, which is high. What matters from our perspective is that government officials have no control over this revenue on a day-to-day basis; nor do they have an ability to interfere with the flow of goods from South Africa. More to the point, domestic producers in the urban areas know this is so, and therefore realize that lobbying policy makers for favors in the trade arena is futile. Absence of an independent trade policy is an extreme form of an "agency of restraint" (Collier 1995).

Delegation of fiscal decision-making power to a non-political body may be yet another solution. The Finance Commission as described by Rao and Singh (2000) in India is one example of such a body. According to Rao and Singh:
The shares of the center and the states and their allocation among different states of both the taxes are to be determined by the Finance Commission appointed by the President of India every five years, or earlier as needed. In addition to tax devolution, the Finance Commission is also required to recommend grants to the states in need of assistance under Article 275.

In practice, however, the Finance Commission has run into a number of problems. Among the problems cited by Rao and Singh are: a lack of clear purpose, a methodology that biases outcomes against poor states, distortion of incentives, a lack of objective criteria for evaluating needs and a lack of an integrated view of the states needs.

We can extend this analysis to consider the case of failed tax reform. How can a government establish credibility to successfully implement a tax reform once a reform has failed? The model itself is silent on this issue but points out some of the potential pitfalls. That is, reform without the institutions in place to make it credible is unlikely to succeed. What are the institutions that make reform credible? These are the same institutions that provide leadership with a long time horizon and the public with avenues for swift and effective retaliation. An example of this type of institution is the post-apartheid constitution for South Africa, through which Nelson Mandela may have succeeded in binding the hands of his successors.

VI. Conclusions

In this paper, we argue that the institutions required to limit government from imposing confiscatory taxes must do two things: they must provide leadership with a long time horizon (e.g., low discount factor) and simultaneously establish avenues for swift and effective retaliation against predatory policies. The political party is one example of such an institution because it provides leadership with an infinite time horizon while, at the same time establishing a body capable of retaliation. When these institutions are not in place and when private investment is irreversible, predatory governments will often succumb to the temptation to tax away sunk investments.

Establishing a well-functioning party system is a slow process, and may not be feasible in the short run. In any case a variety of complementary institutional changes may be needed, including strengthening press freedom and judicial independence. Membership in international bodies with harmonized policies may also help. And production technology itself may change to increase the private sector’s ability to retaliate against taxation.

What is clear from the time-consistency approach is that the specific instruments used for taxation may be less important than whether the government has an incentive for predatory taxation in the first place. Governments that want to impose predatory taxation can do so through a wide array of instruments, so reforms that block the use of one instrument may simply lead to the use of another. But empirically not all situations offer similar incentives for taxation, even within Africa, and we find that observable differences are clearly correlated with both policy regimes and economic outcomes. In this we see cause for optimism, pointing to the possibility of fundamental changes towards more efficient and equitable tax regimes.
Table 8-1. Determinants of Policy and the Impact of Policy on Economic Growth.

<table>
<thead>
<tr>
<th>A. Dependent Variable:</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tax Rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explanatory Variables:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STC</td>
<td>2.56***</td>
<td>0.31</td>
</tr>
<tr>
<td>δ(T,k)</td>
<td>-0.17**</td>
<td>0.07</td>
</tr>
<tr>
<td>π(P^w_k,P^r_k,rd)</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>R^2=0.11, n=551</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Dependent Variable:</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R&amp;D per ag. worker</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explanatory Variables:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STC</td>
<td>-0.06***</td>
<td>0.02</td>
</tr>
<tr>
<td>δ(T,k)</td>
<td>2.19***</td>
<td>0.25</td>
</tr>
<tr>
<td>π(P^w_k,P^r_k,rd)</td>
<td>-0.01</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>R^2=0.29, n=168</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Dependent Variable:</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth 1965-90</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explanatory Variables:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log[taxation]</td>
<td>6.31***</td>
<td>1.74</td>
<td>6.86</td>
</tr>
<tr>
<td>log[rd]</td>
<td>2.35***</td>
<td>0.68</td>
<td>2.55</td>
</tr>
<tr>
<td>log[tax×rd]</td>
<td>1.11***</td>
<td>0.35</td>
<td>1.21</td>
</tr>
<tr>
<td><strong>R^2=0.71, n=19</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D. Dependent Variable:</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth 5 yr. avgs.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explanatory Variables:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log[taxation]</td>
<td>0.151***</td>
<td>0.046</td>
<td>6.86</td>
</tr>
<tr>
<td>log[rd]</td>
<td>0.039***</td>
<td>0.009</td>
<td>2.55</td>
</tr>
<tr>
<td>log[tax×rd]</td>
<td>0.022***</td>
<td>0.007</td>
<td>1.21</td>
</tr>
<tr>
<td><strong>R^2=0.64, n=93</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Coefficients are OLS estimates. Elasticities are computed at variable means. Confidence levels are 99 percent (***) , 95 percent (**) and 90 percent (*). Variable definitions and data sources are as in the text. Control variables not shown in table are a constant for all regressions, initial income and government savings for row (C), and initial income plus country dummies for row (D).
References


Recent World Bank Technical Papers (continued)

No. 463 Stephen Foster, John Chilton, Marcus Moench, Franklin Cardy, and Manuel Schiﬀler, *Groundwater in Rural Development: Facing the Challenges of Supply and Resource Sustainability*

No. 465 Csaba Csaki and Zvi Lerman, eds., *Structural Change in the Farming Sectors in Central and Eastern Europe: Lessons for EU Accession—Second World Bank/FAO Workshop, June 27–29, 1999*

No. 466 Barbara Nunberg, *Ready for Europe: Public Administration Reform and European Union Accession in Central and Eastern Europe*

No. 467 Quentin T. Wodon with contributions from Robert Ayres, Matias Barenstein, Norman Hicks, Kihoon Lee, William Maloney, Pia Peeters, Corinne Siaens, and Shlomo Yitzhaki, *Poverty and Policy in Latin America and the Caribbean*

No. 469 Laurian Unnevehr and Nancy Hirschhorn, *Food Safety Issues in the Developing World*

No. 470 Alberto Valdés, ed., *Agricultural Support Policies in Transition Economies*

No. 471 Brian Pinto, Vladimir Drebentsov, and Alexander Morozov, *Dismantling Russia's Nonpayments System: Creating Conditions for Growth*

No. 472 Jit B. S. Gill, *A Diagnostic Framework for Revenue Administration*

No. 474 Ioannis N. Kessides, ed., *Hungary: A Regulatory and Structural Review of Selected Infrastructure Sectors*

No. 475 Csaba Csaki, Zvi Lerman, and Sergey Sotnikov, *Farm Sector Restructuring in Belarus: Progress and Constraints*

No. 476 Katherine Terrell, *Czech Republic: Labor Market Report*

No. 477 Csaba Csaki, John Nash, Achim Fock, and Holger Kray, *Food and Agriculture in Bulgaria: The Challenge of Preparing for EU Accession*

No. 478 Peter Havlík, *Trade and Cost Competitiveness in the Czech Republic, Hungary, Poland, and Slovenia*

No. 479 Mejmir Mrak, *Communal Infrastructure in Slovenia: Survey of Investment Needs and Policies Aimed at Encouraging Private Sector Participation*

No. 480 Csaba Csaki and Laura Tuck, *Rural Development Strategy: Eastern Europe and Central Asia*

No. 482 Nina Bubnova, *Governance Impact on Private Investment*

No. 487 Tim Schwarz and David Satola, *Telecommunications Legislation in Transitional and Developing Economies*

No. 489 Jesko Hentschel and Radha Seshagiri, *The City Poverty Assessment: A Primer*


No. 491 Tuntivate Voravate, Douglas F. Barnes, and V. Susan Bogach, *Assessing Markets for Renewable Energy in Rural Areas of Northwestern China*

No. 492 Jerry Lebo and Dieter Schelling, *Design and Appraisal of Rural Transport Infrastructure: Ensuring Basic Access for Rural Communities*


No. 494 Gillian Perkins and Ruslan Yemtsov, *Armenia: Restructuring to Sustain Universal General Education*

No. 495 Rogrigo A. Chaves, Susana Sanchez, Saul Schor, and Emil Tesliuc, *Financial Markets, Credit Constraints, and Investment in Rural Romania*

No. 496 Zvi Lerman and Karen Brooks, *Turkmenistan: An Assessment of Leasehold-Based Farm Restructuring*

No. 497 Aldo Baietti, *Private Infrastructure in East Asia: Lessons Learned in the Aftermath of the Crisis*

No. 498 Ali Hashim and Bill Allan, *Treasury Reference Model*

No. 499 Omer Gokcekus, Nick Manning, Ranjana Mukherjee, and Raj Nallari, *Institutional Environment and Public Officials’ Performance in Guyana*

No. 500 Ranjana Mukherjee, Omer Gokcekus, Nick Manning, and Pierre Landell-Mills, *Bangladesh: The Experience and Perceptions of Public Officials*

No. 501 World Bank, *Kosovo: Economic and Social Reforms for Peace and Reconciliation*


No. 503 Geremia Palomba, Milan Vodopivec, *Financing, Efficiency, and Equity in Albanian Education*
