INSIDE

Dixit on Recipes for Development Success

Keeler on Dixit's Recipes

Batten and Szilagyi on Bond Market Development

Umansky and Vegas on School Management Reforms

Gibson and McKenzie on GPS in Household Surveys
The World Bank Research Observer is intended for anyone who has a professional interest in development. Observer articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

The journal welcomes editorial comments and responses, which will be considered for publication to the extent that space permits. On occasion the Observer considers unsolicited contributions. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages to the Editor. Please direct all editorial correspondence to the Editor, The World Bank Research Observer, 1818 H Street, NW, Washington, DC 20433, USA.

The views and interpretations expressed in this journal are those of the authors and do not necessarily represent the views and policies of the World Bank or of its Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of data included in this publication and accepts no responsibility whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.
Evaluating Recipes for Development Success

Avinash Dixit

This article offers a provocative critique of the ability of research on the impact of institutions on growth to offer immediate and practical recommendations for reforming and redesigning institutions in developing countries and transition economies. The literature traces the sources of growth to unalterable historical and geographic features. It contains equally plausible recommendations for opposite courses of action. It is sometimes driven by fads or recommends imitation of the latest success story. Some recommendations are too vague or too general to constitute practical advice. The article suggests a Bayesian diagnostic procedure to identify the causes of economic failure in an individual country as a first step toward remedying the failure. JEL codes: O43, O17, O20, P30, P48.

The main purpose of the most scholarly research, both theoretical and empirical, is to improve our understanding of the phenomena and processes being studied. The work may lead to useful prescriptions for policy, but that is usually the endpoint of a long and winding path. Often, however, especially in matters of economic development and growth, the problems under investigation are large and urgent, so practitioners want immediate answers. Academic researchers are also tempted to offer suggestions for policy, however tentative or incomplete.

Work on institutions and growth is a case in point. When I present my own theoretical models to audiences, I am almost invariably asked to interpret the results for their policy implications. Listeners usually find my responses meager and unsatisfactory, and I am compelled to agree. In this article, I apply the same, admittedly stringent, test of immediate policy relevance to all academic researches on institutions and development. I interpret the findings of this body of work as if they were prescriptions for policy or institutional reform and question the validity and practicality of these interpretations.
I am deliberately provocative and critical, but try to be evenhandedly so. I hope
to give everyone some incentives to think further and harder. I also hope to help
scholarly researchers better see their own work in the context of the bigger
picture and help practitioners better appreciate the difficulties of drawing impli-
cations for action from an ongoing academic exchange.

My primary focus is on institutions of economic governance—protection of
property rights, enforcement of voluntary contracts, and provision of public infra-
structure and services that support private economic activity. This relates to the
broad question of the emergence and reform of deep institutions such as democ-
archy, insofar as that influences the shape and function of the institutions of eco-
nomic governance and in turn affects economic performance. It also relates to
some specific organizations that implement the rules of the game established by
the institutional setting and even to some specific policies that illustrate the issues
involved in thinking about institutional change. There is a spectrum from deep
institutions to specific policies: my discussion is located somewhere in the middle,
with the focus blurring toward each end. One more restriction, dictated by the
limitations of my own expertise, is that I focus on microeconomic issues and insti-
tutions, not those of macroeconomic policy.

Even within these confines the literature is huge, and I cannot hope to include
even a substantial fraction of it in the space available. However, as the policy
recipes that emerge from this literature are almost invariably unsatisfactory, being
omitted from the list and the implied criticism, however friendly and constructive,
may perhaps be regarded with relief.

Critical assessments of the contribution of academic research to development
policy have a long tradition. Most recently, Easterly (2001) has produced a com-
prehensive, compelling, and entertaining the critique of theories of economic
growth. He concludes that the prescriptions derived from these theories—policies
to increase saving, investment, foreign aid, and education; to modernize technol-
ogy; and to reduce population growth and debt—failed mostly because other pol-
icies of the developing country governments often maintained or created perverse
incentives. He concludes that the path to success must begin by setting up
“quality institutions” (p. 252), but he says little about how to do this. I will try to
extend the critique to the research on institutions by subjecting this literature to a
prospective test of usefulness: a retrospective test will have to await the outcomes
of any attempts to use it.

Here are some examples of broad categories of ways in which the accumulated
research on the role of institutions in development stops short of giving any
useful or reliable policy prescriptions:

- Some of the works find that development success is contingent on some his-
torical or geographic preconditions that most countries do not meet.
On many crucial issues, from deep institutional divides of democracy or authoritarianism, to general policy approaches such as comprehensive change or gradualism and sequencing of reforms, to specific policies toward population, inflation, regulation, and so on, the literature includes support for diametrically opposed prescriptions, each seeming quite plausible when presented on its own. Thus, unlike the theories of economic growth, which Easterly criticizes for serially offering definitive answers that did not work in practice, theories of institutions simultaneously offer different and conflicting answers.

- The recipes for success are often motivated by ideological fashions or a herding instinct that follows the latest perceived success.
- The recipes that seem most promising are often at a level of generality that stops far short of practicability. Examples are advice to build institutions to complement existing ones or to build participatory democracy; this leaves out all details of how to identify the complementarities or how to convert an authoritarian regime into a participatory democracy.

My conclusion from this overview of the literature is that the research has not elucidated the mysteries of successful development, and we cannot draw many reliable and practicable policy conclusions from it. But I do not want to be entirely negative. In the concluding section, I suggest a framework or methodology of research that combines general conceptual and empirical findings from academic research and the experience of practitioners to help narrow or identify the causes of failures in individual countries. I hope that it can serve as the first step toward remedying the failures or removing the obstacles.

**Infeasibility—The “Irish” Recipes for Success**

Numerous researchers have found significant correlations across countries between various geological, ecological, geographic, and historical data and economic performance. Taken literally, these findings constitute a message of pessimistic determinism: if your country lacks the right prior or starting conditions, its economic future is bleak. Of course, many of the correlations are not to be taken literally, but I will start by listing them all on an equal footing of literal interpretation and then develop further arguments and interpretations. To make them amusing and memorable, I call these findings “Irish” recipes for success, after the story of the Irishman who, when asked for directions, replied “If I wanted to go there, I wouldn’t start here.”

Many geological and geographic characteristics have the appearance of determinism. The fixed or very long-run processes described by Diamond (1997) are
the most extreme examples of such determinants of economic outcomes: an east–west rather than a north–south alignment, the right kinds of native animals and plants, and so on. We are also told that it helps to be far from the equator, in the temperate rather than the tropical belt (Gallup, Sachs, and Mellinger 1999; Hall and Jones 1999), unless your country is a small Pacific island, in which case distance from the equator is conducive to a total collapse of your ecology, economy, civilization, and population (Diamond 2004). Being far from the core areas of the world economy is a major piece of bad luck, and being landlocked makes it worse (Gallup, Sachs, and Mellinger 1999). You might have thought that your country was fortunate to be endowed with abundant natural resources, but this is debatable. A considerable body of literature argues that resource abundance worsens your prospects. Sachs and Warner (2001) summarize and extend these findings; see also Engerman and Sokoloff (1997).1 The problem is not confined to mineral resources: if your land is suitable for plantation agriculture, that can also be bad for your economic performance (Engerman and Sokoloff 1997).

On the other side of the balance, if your region was sparsely populated and poor at the time of the spread of European colonialism in the 1500s, your country is more likely to be economically successful now (Acemoglu, Johnson, and Robinson 2002; Engerman and Sokoloff 2002). It was a special help to be colonized by the British, so that a large fraction of your population became fluent in English (Hall and Jones 1999) and inherited British rather than French or other European legal and administrative institutions, giving better protection to the rights of small shareholders and therefore allowing broader and more efficient capital markets to develop (La Porta and others 1998, 1999).2 If local diseases were more likely to take a heavy toll on the early colonizers, your economic performance today is less likely to be good (Acemoglu, Johnson, and Robinson 2001). And numerous economic and social scientists from Weber onward have appealed to higher authority, finding roots of success in Protestantism, Confucianism, and so on.

Interpreted literally as recipes or policy recommendations, these require a developing country to use plate tectonics to move itself to a more favorable location or to turn the clock back and invite British colonizers. Of course cleaning up the local disease environment and getting rid of mineral resources beforehand. As a practical matter, these findings are merely telling countries to accept their fate. At best, they might be reinterpreted as telling the fortunate countries to share their good luck by giving massive amounts of aid to the unfortunate.

Of course, this is not the interpretation the researchers intend: they intend many of their history and even geography variables to have only indirect effects on economic outcomes through some other proximate determinant of success or to be mere econometric instruments used for identifying the direction of
causation. This is emphasized by the authors of these articles and by others such as Rodrik (2004a, p. 4), who adds cautiously: “Finding an appropriate econometric instrument is not the same as finding an adequate explanation.” The supposed proximate determinants of success are categorized as institutions (Engerman and Sokoloff 1997, 2002; Acemoglu, Johnson, and Robinson 2001, 2002) or social infrastructure (Hall and Jones 1999). Here are some examples of the suggested channels.

If a colony had rich mineral resources or climate and soil conditions conducive to plantation agriculture, and a large or dense population that could be press-ganged into work in such activities, then the European colonizers established institutions such as slavery and inequality that facilitated extractive and plantation activities. Institutional change is a very slow process; old institutions persist or shape today’s institutions. But they are less suitable for modern economic conditions, where broader participation in decision making, greater access to education and economic opportunities, and protection of everyone’s property rights are needed for success. Areas that were sparsely populated and poor in the 1500s were likely to lack such get-rich-quick opportunities. There, the European colonizers set up institutions conducive to longer-term economic success. (See Engerman and Sokoloff 1997, 2002; Acemoglu, Johnson, and Robinson 2002. See also an excellent account of this research in Hoff 2003.)

If European colonizers contracted local diseases and experienced high mortality rates, they were likely to establish extractive economies and institutions. If they had good prospects of long lives as producers, they were likely to settle and establish economic activities and supportive institutions conducive to long-run economic success. (See Acemoglu, Johnson, and Robinson 2001; see also the critique by Albouy 2004 and the reply by Acemoglu, Johnson, and Robinson 2005.) More broadly, Hall and Jones (1999, p. 100) regard “various correlates of the extent of Western European influence,” primarily distance from the equator and the prevalence of English, French, German, Portuguese, and Spanish as first languages in the countries, as instruments for “social infrastructure,” measured by an index of lack of corruption (government anti-diversion policies) and openness to international trade.

The need for instruments arises because of potential reverse causation: good institutions are likely to be conducive to good economic performance, but citizens of economically successful countries may also demand and implement good institutions. However, Kaufmann and Kraay (2002) find a weak but negative reverse causation, suggesting the absence of a virtuous circle between better governance and better economic outcomes. Keefer (2004a) also argues for a weak or even negative reverse causation, combining conceptual arguments and empirical evidence. But even negative reverse causation can create an econometric problem requiring instrumental variables.
Formally, let $y_i$ denote the economic performance of country $i$, $x_i$ be a vector of measures of its institutions, and $z_i$ a vector of other relevant variables. The hypothesis being tested is

$$y_i = a_1 + a_2 x_i + a_3 z_i + e_i.$$  

There is a reverse causation from the $y_i$ to the $x_i$, so one invokes instrumental variables $g_i$ that are assumed to affect the $x_i$, but have no direct effect on the $y_i$. Thus

$$x_i = b_1 + b_2 y_i + b_3 z_i + b_4 g_i + u_i.$$  

Equation (1) can then be estimated by two-stage least squares, estimating the $x_i$ from the first-stage equation

$$x_i = c_1 + c_2 z_i + c_3 g_i + v_i,$$

and using the estimated values in the regression for equation (1). This requires that the instrumental variables $g_i$ do not appear directly in equation (1) and are uncorrelated with the error term $e_i$ in that equation. Various stories can then be constructed for why that should be so. For example, Hall and Jones (1999, p. 101) comment about their “European influence” instruments:

we must ask whether European influence was somehow more intensively targeted toward regions of the world that are more likely to have high output per worker today. In fact, this does not seem to be the case. On the one hand, Europeans did seek to conquer and exploit areas of the world that were rich in natural resources such as gold and silver or that could provide valuable trade in commodities such as sugar and molasses. There is no tendency today for these areas to have high output per worker. On the other hand, European influence was much stronger in areas of the world that were sparsely settled at the beginning of the sixteenth century... Presumably these regions were sparsely settled at that time because the land was not especially productive given the technologies of the fifteenth century. For these reasons, it seems reasonable to assume that our measures of the Western European influence are uncorrelated with [the error term].

This discussion also relates to the instruments and their discussion in the other work of this kind cited above.

Many of these stories sound plausible. But how does one formally test them? In other words, how does one test the difference between the system of equations (1) and (2) and an equation like this:

$$y_i = d_1 + d_2 z_i + d_3 g_i + w_i.$$
that allows economic success to depend directly on the supposed instruments and not on
the supposed proximate institutional determinants?

If there is just one instrument, for example, colonial mortality in the primary
specification of Acemoglu, Johnson, and Robinson (2001), then the system of
equations (1) and (2) is exactly identified and econometrically indistinguishable
from equations (3) and (4), unless one is willing to make other assumptions
about error variances or measurement errors. This is the procedure adopted by
Hall and Jones (1999), Acemoglu, Johnson, and Robinson (2001), and Kaufman
and Kraay (2002). It involves some other untested assumption. If there is more
than one instrument ($g_i$ is a vector), one can perform over-identification tests, but
these are weak and too likely to accept the null hypothesis that the $g_i$ terms are
excluded from equation (1). In Acemoglu, Johnson, and Robinson (2002), the
geographic reversal of fortune instrument is tested in such a way, assuming that
mortality among colonists is a valid instrument. But that is not problem-free, as
was just argued.

My view is that the notion that geographic and historical variables are merely
instruments for institutional determinants of economic success is supported more
by the intuitive appeal of the stories told than by the statistical significance of the
tests performed. The value of rhetoric should not be ignored, but I wish the
econometric evidence were more compelling. More and sharper research is badly
needed.

But what policy implications follow if we proceed as though we were convinced
by these stories and accept that institutions are important? The literature
reviewed above gives us no guidance. Whether geography or history have a direct
effect or an effect through institutions, the recommendation to change one's
geography or history is useless. We have to forget about history and geography
and try to affect the relevant institutions directly. For advice on that, we must
turn to some other research.

Contradictions—“Doctor Dolittle” Recipe Pairs

There is no shortage of academic research that identifies institutions and policies
to promote economic growth and development and proffers advice on how to
acquire good institutions. Indeed, Rodrik (2000) titles his paper “Institutions for
High-Quality Growth: What They Are and How to Acquire Them.” This literature
presents some convincing evidence and arguments that institutions that offer
credible commitments to protect property rights and to enforce contracts are
good. But there is much disagreement about which institutions are better for this
purpose. For every paper that endorses one kind of institution or policy, there is
another that claims the opposite. Each is written by a prominent economist and
contains impressive arguments and evidence to support its recommendation.
The following sections examine several dimensions of the contradictory conclusions of the literature on institutions and policies. They range from broad choices about the deepest institutions that govern society and polity, to minutiae on specific policies. Once again I have chosen a whimsical and memorable label for this literature. To remind us of the two-headed animal “pushme-pullyou” in the book and musical, I call these the “Doctor Dolittle” pairs of recipes for development.

**Democracy or Authoritarianism?**

Many cross-country regressions examine whether the democratic or authoritarian governments are better for growth and find mixed results. For example, Barro (1999, p. 61), who finds a relatively poor fit and an inverse U-shaped relationship, suggests that “more democracy raises growth when political freedoms are weak, but depresses growth when a moderate amount of freedom is already established.” Persson (2005, p. 22), using cross-sectional as well as panel data, finds that the crude distinction between democratic and nondemocratic forms of government is not enough; the precise form of democracy matters for policy design and economic outcomes. “Reforms of authoritarian regimes into parliamentary, proportional, and permanent democracies seem to foster the adoption of more growth-promoting structural policies, whereas reforms into presidential, majoritarian, and temporary democracy do not.” However, Keefer (2004b, p. 256), after surveying a wide-ranging literature on electoral rules and legislative organizations, concludes that they affect policies but are not a crucial determinant of success: “electoral rules ... almost surely do not explain why some countries grow and others do not,” and “the mere fact that developing countries are more likely to have presidential forms of government is unlikely to be a key factor to explain slow development.”

What can explain the claimed superiority of democracy? Rodrik (2000) emphasizes the importance of local knowledge for successful institution building and argues that participatory democracy is a meta-institution that facilitates such use of local knowledge and thereby enables higher quality growth. Besley, Persson, and Strum (2005), using data from the United States, argue that political competition is a key to better economic policies and outcomes, and this also has intuitive appeal. Besley and Burgess (2002), using panel data from India, find that an informed and active electorate leads to effective incentives for governments to respond to economic problems and that mass media play an important part. Democracy succeeds by facilitating voice and participation. Bardhan (2005) also stresses the importance of democratic participation in a generalized interpretation of the “rule of law.” Islam (2003) emphasizes the role of information and transparency in improving the quality of governance institutions.
Some who favor democracy over authoritarianism in a comparison of ongoing systems may nevertheless be concerned about the costs of disruption in a transition to democracy. For them, Rodrik and Wacziarg (2005) bring good news. Using annual panel data for more than 150 countries (with separate subsamples analyzed for some questions), they find that transitions from authoritarian to democratic regimes lead to striking and statistically significant improvements in GDP growth per capita: for example, the effect is as large as 2.8 percent a year for Sub-Saharan Africa. Growth also becomes less volatile after a transition to democracy. Declines in growth precede, not follow, such transitions.

These are just a few examples from a vast literature, and they add up to a message that is pleasing and even uplifting to many modern academics and policy practitioners: democracy is good not only for its moral and human appeal, but also for its economic performance. This would be a "warm glow" recipe or a "happy ending."

But there is an equally impressive emerging literature that makes a serious case for authoritarian governments and institutions for starting growth and development. Glaeser and others (2004) argue that the developing countries that achieve economic success do so by pursuing good policies, often under dictatorships, and only then do they democratize. While these conclusions are controversial, these authors' criticisms of the measures of institutions used in the research that argues for the primacy of institutions in general, and of democracy in particular, are telling. Giavazzi and Tabellini (2005) find a positive feedback between economic and political reform, but they also find that the sequence of reforms matters, and countries that implement economic liberalization first and then democratize do much better in most dimensions than those that follow the opposite route.

This debate could go on for a long time with contending arguments about basic concepts and theories as well as empirical methodologies, data, and anecdotes. Here I would like to say something brief about two questions.

First a conceptual matter: what policy feature or features are important for good economic outcomes, regardless of the kind of government that makes those policies? There is broad agreement in much of the literature that the credibility of commitments is vital. Acemoglu (2003) argues that the lack of third-party enforcement in political contracts makes it harder to make credible commitments, and that this explains the absence of a Coase theorem ensuring efficient outcomes in political bargaining. In a comprehensive overview of the dimensions of governance and their effects on economic development, Keefer (2004a, b) also identifies the ability to make credible commitments as crucial. Bardhan (2005, ch. 4) offers a good discussion of the concept of credibility and of its lack as a barrier to reform. In an illuminating analytical narrative of Argentine politics, Spiller and Tommasi (2003, p. 281) argue that the institutions and organizations of political
bargaining in Argentina, and the volatility of its economic environment, combined to inhibit the capacity of its policymakers to "undertake efficient inter-temporal political exchanges," and this led to "low-quality policies" on many issues such as regulation and control of inflation.

How do democracy and authoritarianism compare in this respect? Keefer (2004b) stresses the role of checks and balances in achieving credible commitments, and these are more likely to be present and effective in democracies. However, one might argue that commitments from authoritarian rulers should be more credible, if anything, so long as they are stable in their rule. And such rulers do have incentives to pursue policies that enhance economic success. Even if their aims are predatory, as long as their positions are sufficiently stable, they will achieve the largest increase in their own take if the pie is larger—this is Olson’s (1993) "stationary bandit." If they fear being overthrown, improving the people’s economic well-being may be the best way of postponing that fate.

China is usually cited as an authoritarian regime in which property rights are not formally protected but are de facto quite secure to the point that the country attracts large investments from abroad. The literature offers various explanations. Qian (2003) and Rodrik (2004b) attribute this to the Chinese institutional innovation of township and village enterprises, which "force[ed] entrepreneurs into partnership with their most likely expropriators, the local state authorities. . . . Local governments were keen to ensure the prosperity of these enterprises as their equity stake generated revenues directly for them" (Rodrik 2004b, p. 11). This puts the local authorities in a role similar to Olson’s stationary bandit. However, the strategy sounds similar to the concept of "insider privatization" pursued in Russia and supported by Shleifer and Treisman (2000, pp. 31–2), which did not work so well there. What was the difference?

McMillan (2003, pp. 98–100) argues that in the phase of China’s agricultural reform, "productivity gains were achieved without formal legal recognition of farmers’ ownership rights. . . . Although the authorities are able to renge on contracts, they have refrained from doing so with sufficient predictability that the farmers are motivated to be productive." He finds the probable reason in the "specifics of time and place. . . . The communist government faced no challenge, [but] its legitimacy as the government, and its ability to preempt any future political opposition, rested on its delivering economic growth. High officials in Deng Xiaoping’s government understood enough about economics to recognize that the growth requires markets and markets require assured property rights. The Communist Party had retained its highly disciplined organization and so was able to prevent self-seeking behavior by low-level officials." This is the idea of a dictator operating under a constraint that captures his fear of losing power. Indeed, China appears to retain several aspects of democracy at lower and middle levels of institutions and economic policy-making: there are some genuinely contested
elections at these levels, press criticism of officials at this level is tolerated and perhaps even encouraged, and corruption is swiftly and severely punished when detected. Only at the top level is the Communist Party's rule rigid and unchallengeable. Thus, it may be possible to combine the two forms of government institutions for optimal economic performance.

However, there is always the problem of how to ensure up front that your authoritarian ruler turns out to be a benevolent promoter of economic success and not a predatory despot. And if your country has the misfortune of getting a predatory despot, you face the bigger problem of how to get rid of him. Similarly, the voluminous literature has almost unanimously made a convincing case that the credible protection of property rights is essential for investment and innovation and is therefore a necessary condition for development success. But this literature provides no guidance on the operational question of what can be done if despots cancel these rights on a whim or democratic governments do so at the behest of politically important special interests.

**Formal or Informal Governance Institutions?**

Institutions to protect property rights and enforce contracts can be formal (the state's apparatus of legislation, policing, and the judiciary) or informal (social networks, communication channels, and norms) and for-profit information and enforcement services, including organized crime. Many variants and mixtures of formal and informal systems exist in most countries. In some countries, systems of private adjudication are supported by formal enforcement: for example, industry-based arbitration panels can be recognized by courts, which enforce the judgments issued by arbitrators. Rubin (1994) recommends this approach for building governance institutions in transition economies. In some countries, private enforcement of formal laws, ranging from private initiation of litigation to private infliction of punishment, is also practiced. Berglöf and Claessens (2004) and Hay and Shleifer (1998) favor forms of this approach.

What are the relative merits or flaws of formal and informal institutions of economic governance? The literature contains an often bewildering and mutually contradictory spectrum of arguments and evidence. A useful conceptual framework for studying formal–informal variants is the distinction between relation-based and rule-based governance of contracts. In relation-based governance, transactions occur between parties who are both members of the same group or network based on such ties as language, ethnicity, and locality. Information circulates within the network and helps each member in searching for a partner with which to transact and in obtaining information about the capability and trustworthiness of prospective partners. Contracts are personal and implicit agreements. Any breaches of promise are communicated to
the group, and other members of the group can punish offenders by refusing to trade with them or by imposing extra-economic sanctions such as social ostracism or worse. Rule-based governance relies on formal contracts and their enforcement by the courts or the police, if necessary.

Although this distinction was known and studied for many years in many contexts, Li (2003) gave it sharper analytic content by clarifying the main difference in the costs of establishing and operating the two systems. Relation-based governance has few fixed costs. It does not require a framework of laws, regulations, or courts; one just starts dealing with one’s friends and neighbors. But its marginal costs are substantial and increase as the scope of trade expands. One deals initially with close neighbors, whose trustworthiness is better known and who are more likely to behave well than are more distant acquaintances or strangers because future links among close neighbors are likely to be larger and more frequent. But as one’s business expands, it becomes necessary to deal with those more distant acquaintances or even strangers, who are less likely to be trustworthy. In contrast, rule-based governance has high-fixed costs of legislation, regulation (such as reporting and auditing requirements that create information about traders), and enforcement. But once these are in place, one can deal with strangers at low-marginal cost. Because of these cost differences, the relation-based system is better at small scales of transaction, and the rule-based system at large scales. Dixit’s (2004) formal model examines and compares equilibria under the two systems and elucidates the interaction of different parameters and mechanisms that determine the limits of the relation-based system.

This was precisely what Greif (1993, 1994, 1997) found when comparing two groups of traders around the Mediterranean in the medieval period. The Maghribis relied on relation-based governance within their tight group of Jewish traders; this worked well up to a point but eventually could not cope with the expansion of trade to newer and more distant locations. The Genoese traders used a more formal system of registering contracts with the authorities in Genoa, who investigated and adjudicated any disputes. This system proved better as trade expanded. More recently, microcredit has succeeded in many small communities and groups, but expanding its scope to larger financial markets for lending, borrowing, and investing has proved problematic. A possible approach is to link the small communities together at an upper tier using the kind of system Bernstein (1992) describes for the diamond industry or the “community responsibility system” in medieval Europe studied by Greif (2004).

The worsening quality of information and communication is not the only problem that limits the expansion of informal institutions. Another issue is inherent in the concept of a “community” that develops in these informal institutions. This is a group of “insiders,” which by its nature requires a concomitant definition of the “outsiders” who are to be excluded. Therefore, informal
institutions usually erect entry barriers that inhibit fuller exploitation of opportunities and development of capital markets. Fafchamps (2004, pp. 455–56) develops this idea and offers examples.

When comparing the relative merits of formal and informal or rule-based and relation-based institutions of governance, it may be important to distinguish between enforcement of contracts and protection of property rights. Informal and relation-based systems seem to be less successful in protecting property rights than in enforcing contracts. Li (2003, p. 657) confines most of his discussion to contracts and tacitly assumes a background of formal institutions that include property right protection: he says that relation-based governance “requires only minimum public order—that is, the general absence of rampant robberies or confiscation” to get going. And Gambetta’s (1993, p. 198) study of the Sicilian mafia indicates that their performance in enforcing contracts could be socially beneficial, whereas in their role of protecting property, “protectors, once enlisted, invariably overstay their welcome.” Even with contracts, information services about trustworthiness do not generate problems of violence, whereas enforcement services to punish breaches do (Gambetta 1993; see also Dixit (2004)).

De Soto (2000, pp. 5–6) emphasizes the importance of formal recognition of property rights:

in Asia, Africa, the Middle East, and Latin America, . . . most of the poor already possess the assets they need to make a success of capitalism. . . . But they hold these resources in defective forms: houses built on land whose ownership rights are not adequately recorded, unincorporated businesses with undefined liability, industries located where financiers and investors cannot see them. Because the rights to these possessions are not adequately documented, these assets cannot readily be turned into capital, cannot be traded outside of narrow local circles where people know and trust each other, cannot be used as collateral for a loan, and cannot be used as a share against an investment.

But he recognizes the importance of building on local knowledge (De Soto, 2004, p. 108): “As I strolled through rice fields [in Bali], I had no idea where the property boundaries were. But the dogs knew. Every time I crossed from one farm to another, a different dog barked. Those Indonesian dogs may have been ignorant of the formal law, but they were positive about which assets their masters controlled.” So officials who wanted to set up a formal property system could “by traveling their city streets and countryside and listening to the barking dogs, . . . gradually work upward.”

Most theoretical and empirical studies find that the purely informal institutions of governance eventually act as a constraint on growth and expansion of trade. No country switches entirely to purely formal institutions; even in the most
advanced countries much economic activity continues to be governed by relational and private ordering "under the shadow of the law." Bernstein (1992, 2001) offers some examples of such industry-based adjudication in the United States. But it remains important for developing countries and former socialist countries to develop more formal institutions that can govern increasing volumes of arm's length transactions.

This transition brings its own difficulties, and evidence as well as arguments on this issue yields mixed results. Some formal institutions can meld well with the existing informal ones. Rubin (1994) and Dixit (2004) argue that arbitration does, although Widner (2000) has challenged the efficacy of specialized courts.

In other instances, the two types of institutions clash, and the formal one may lose. Ensminger (1997) gives an example of land title reform in Kenya. Traditional land rights in many parts of Africa are a complex system. Clan chiefs grant titles to individual families. Sales are subject to their approval and also to that of family heirs since all sons usually have expectations of equal division. Many family stakeholders have usufruct rights. When the Kenyan government attempted to impose a system of formal land titles, this ran into conflict with the traditional arrangements. The expected capital market did not develop because lenders knew that foreclosure was infeasible in the face of opposition from family and community, so the land could not be used as collateral. Attempts to consolidate scattered holdings for scale economy reasons did not work because there was a good economic reason (insurance) for the scattering. Many formally registered titles are now being allowed to lapse and revert to older arrangements, and the laws are being changed to more closely resemble traditional forms of ownership. Finally, the theoretical literature, using a repeated-game framework, shows how a partial improvement of an imperfect formal system, by providing a better outside alternative and thereby lessening the harmful consequences of breaking a relational contract, can worsen the outcomes of the informal system (Baker, Gibbons, and Murphy 1994; Dixit 2004).

Linked to the conflicting arguments and evidence about formal and informal institutions is the related question of whether to use the readily available models of Western legislative and judicial institutions or to develop ones specifically tailored to each country. Many of the pertinent issues have already been touched on here or will be in the following section, so they are not discussed again except to mention their bearing on this question.

**Comprehensive and Rapid or Sequential and Gradual Reforms?**

The rapid or gradual reform debate is long-standing, and it continues unabated. There are four main arguments for speed. First, opportunities for reform present
themselves rarely in political negotiation, so they should be seized when they do arise. Second, reforms cause some pain to some people or groups, and gradualism prolongs the pain, risking the whole reform. Third, policy actions convey useful signals about the government’s intentions to financial markets and investors, and in the standard Spencian manner a reforming government must engage in excessive signaling to separate itself from a less purposeful one. Fourth, the dimensions of reform are strategic complements, and results will be poor unless all dimensions are tackled jointly (see Kremer 1993a for an “O-ring” theory of this general class of phenomena).

Others favor a more cautious approach, presenting two main arguments. First, drastic reforms, by inflicting severe pain, are likely to arouse strong opposition and thereby create a greater risk of political opposition, leading to failure. Second, institutional change is a slow process since it requires changes in long-held expectations.

Prominent early supporters of speed and comprehensiveness include Åslund (1995); those on the side of caution include Murrell (1992) and Desai (1995). The World Bank (1996), in World Development Report 1996: From Plan to Market, emerged in favor of speed but with many cautions and caveats. Heybey and Murrell (1999) offer an empirical assessment based on more recent data. They find that a country’s initial conditions are more important than policy changes in determining its economic performance during the first few years of transition; that is, whether the reforms are rapid or gradual is less important. They also find a negative feedback in reform: higher levels of initial liberalization are likely to slow the subsequent reforms.

A somewhat different recent argument in favor of one kind of gradualism, and a new view of optimal sequencing, says that the most productive institutional changes in the early phases of development are small and easy to implement. The harder job is to build on them and sustain growth beyond a medium level. Thus Pritchett (2003, p. 148) concludes from an analytical case study: “Under a regime that has reasonable institutional stability and is not completely dysfunctional, a rapidly increasing level of GDP per capita is possible up to semi-industrialization. . .a rough and ready system for enforcing contracts and providing for stability of investors’ expectations” suffices to “support quite rapid growth if beginning from a low base.” Even some corruption can be tolerated so long as it is “predictable,” meaning that investors are clear about whom to bribe and how much to pay them and can rely on the official delivering on the quid pro quo. “However, what trips countries up is the transition from one set of institutions to another [emphasis in the original]” because this requires a change in expectations and creates uncertainty during the transition.

Rodrik (2003, p. 17), in his overview of Pritchett’s paper and several other others, similarly concludes that “the policies required to initiate a transition from
a low-income equilibrium to a state of rapid growth may be qualitatively different from those required to reignite growth for a middle-income country." This fits with the ideas of Li (2003) and Dixit (2004) discussed above: so long as some basic protection of property rights exists, even formal contract enforcement may not be needed in a small economy—self-enforcing, relation-based governance can work well. But to go beyond that stage, formal institutions of laws, regulatory agencies, courts, and the police are needed, and a malfunction in any part of this complex can create serious obstacles and setbacks to growth.

This argument can be linked to the idea of strategic substitutes and complements. Where different institutions are substitutes, albeit imperfect ones, it is fine to start with the simpler one and then move gradually to a more complex and better one. But if the institutions are complementary, they need to be implemented in a package. For example, privatization may fail unless there is an adequate structure of regulation, and a good forum for adjudicating commercial disputes may be futile without a way, public or private, to enforce its decisions. Therefore, for gradualism to succeed, a policy practitioner must know which policies or institutions are substitutes and which are complements. This is a difficult task, and the literature offers only a few, incomplete hints.

Crises—Good or Bad?

Also linked to the speed and comprehensiveness of reforms is the possibility of beneficial consequences of crises. Olson (1982) presents this thesis in the clearest form. He argues that many inefficiencies in prevailing institutions arise to serve organized special interests, and the entrenched political power of these groups prevents reforms that would be beneficial in the aggregate. Crises weaken or dissolve these special interests, and allow a fresh start. Williamson (2000) makes some related observations. And Calvo (2005) points out that many recent crises have been followed by a sharp recovery in which such usual factors as investment, domestic bank credit, and the current account deficit were only small contributors—a miraculous Phoenix-like rise from ashes.

It is hard to imagine a country deliberately risking a grave crisis in order to start the Olsonian process. Moreover, other researchers argue that crises are not necessarily beneficial. The theoretical model of Hsieh (2000) clarifies the issues. The dynamic negotiation on policy reforms following a crisis is a war of attrition or a game of "chicken," the question being who will concede first. With incomplete information about the alternatives, value systems, and impatience of other players, such games can have equilibrium strategies that involve lengthy periods without agreement, and may sometimes result in total breakdowns. Whether crises change these equilibria for the better is ambiguous. Diamond (2004) presents some dramatic practical examples of crises leading to total collapse.
Population—Resource or Curse?

Many specific issues have received similar treatment in the literature as institutions and their reform, with conflicting results and prescriptions. Population is a good example. A general presumption dating back to Malthus is that a large and rapidly growing population is bad for economic success. In the standard neoclassical growth model, a larger rate of population growth leads to a steady state with lower output per capita. When natural resource depletion is also considered, the consequences can be more dire. Diamond (2004) is only one of the latest of a long line of prominent scientists who have sounded alarms on this issue.

But again, there are strong counterarguments. Boserup (1981) argues that high-population density is conducive to faster technical progress. And growth models that incorporate such positive feedbacks, for example, Lee (1988) and Kremer (1993b), find superexponential growth. Platteau (2000) finds that high density is also conducive to the emergence of better networks and institutions.

Whom to Believe, and What to Do

I have sketched pairs of starkly conflicting research findings on several important issues of institutions and policy. This can leave anyone who is not an expert in a particular area in a state of confusion and indecision. That might not be a bad outcome. Confronted with such uncertainty about the true model and the factual details, practitioners who judge the balance of considerations to favor one side would nonetheless be wise to proceed with caution. Rarely are the arguments and evidence from the other side clearly wrong or unambiguously inferior. The decisions being made have immense consequences, and they are difficult to reverse. Therefore, any uncertainty weighs heavily in the decision. The favored choice must not merely be better than the other, but better by a sufficiently large margin to justify going ahead. Otherwise, waiting for better information is optimal (Dixit 1992; Dixit and Pindyck 1994).

In some circumstances, there is another problem when it comes to action. Even if you are convinced by one side in the argument and want to follow through into institutional reform or policy action, there may not be much you or almost anyone else can do. Suppose you are the minister for the economy in an authoritarian regime. You read the writings of Rodrik and others about the virtues of democracy and are thrilled by the thought of having this “meta-institution” that harnesses “local knowledge” in your country. What do you do? (Of course, if you are the minister for the army in a democracy and are convinced by the pro-authoritarianism arguments, you may find it easier to implement your favored institutional reform!)
Shifts of Consensus

In the previous section, I laid out several stark disagreements among researchers and showed how difficult it is to sort out reliable and feasible policy recommendations from this chaos. This “cross-sectional problem” is augmented by a “time-series problem.” Among researchers as well as practitioners there may be a reasonable consensus at any one time, but the consensus can shift dramatically over time. Here is a brief outline of some such changes and the warning they should convey.

Chasing the Latest Success Story

At any time some country is doing well, and academic as well as practical observers are tempted to generalize from its choices and recommend the same to all countries. After a decade or two this country ceases to do so well, and some other country using some other policies starts to do well and becomes the new star that all countries are supposed to follow.

In the 1950s, many intellectuals expected the Soviet Union to grow rapidly, and some form of central planning, including import-substitution and state investment in heavy industry, was offered as the key to success. In the 1970s through the 1990s, Japan and other emerging East Asian economies became the heroes, and markets and openness became the vogue. Now China is the great example to follow. What next?

Role of International Organizations

For the last 50 years or longer the World Bank, the International Monetary Fund, and other international organizations have given advice, finance, and aid to developing countries and transition economies. But perceptions are shifting about exactly what their role should be. At one time they were supposed to be the enforcers and guarantors of a country’s commitment to policy reform: conditionality was the key. But the conditions were criticized as one-size-fits-all recommendations. Designing different policies for each country based on local knowledge to fit its conditions was argued to be vital, and countries’ “ownership” of reforms became the buzzword.

The idea that international institutions can act as facilitators or catalysts for institutional change or policy reform (McCarthy, Bader, and Pleskovic 2003) seems attractive; it can combine the best features of discipline and ownership. Dollar and Levin (2005) present evidence that the World Bank’s project finance and aid are most productive in countries with good institutional quality; thus, countries and international institutions appear to be strategic complements in
producing economic success. But we have little hard evidence on whether local ownership translates into better—more credible, less corrupt—institutions. All of the issues mentioned above in the context of authoritarianism or democracy can arise equally at the local level.

The Napoleon Prescription

Faced with all these contradictions and shifts, I can identify only one consistently valid policy prescription. It is the quality Napoleon valued most in his generals—luck. Researchers want to identify causes and practitioners want to know what they can choose and change. Both sides may have neglected the important role that luck has played in many countries' development successes or failures. Easterly (2001) is rare among economists in giving luck a substantial role and discussing it in considerable detail. I offer a few supplementary examples.

Japan's success with the auto industry owes much to luck. The improvement in the quality of their small cars came at just the right time in the mid- to late-1970s when rising gas prices had shifted US demand toward small cars and the US industry had not shifted its designs and production to match. And the Japanese quality upgrading in the early 1980s also hit it just right, as the "voluntary" export restraints imposed by the United States on Japan added an equal absolute shadow value of the quota to the prices of all cars and thereby changed the relative price ratio in favor of larger cars. As for bad luck, think of the many instances where large investments of resources and time in the building of tourism in a country go to waste because of some natural disaster or political turbulence.

Division of Labor and Diagnostics

Why do we find so many contradictions in the research on the determinants of economic growth and development and so many corresponding conflicts in the recipes for development success? Some explanations lie in the diversity of theories and perspectives that economists and academics from other fields bring to this inquiry: development economics is undergoing a paradigm shift, from theories that view resource and technology constraints as key obstacles to growth, to theories that view information asymmetries as key constraints to the operation of markets, and now to theories that view institutions as keys to success or failure. In the process of a paradigm shift, no one's point of view will dominate, and major differences will persist for a while. This is not very satisfactory for practitioners, who cannot afford to wait until the dust settles and a new paradigm emerges.
Another part of the explanation for the current state of differences lies in the diversity of data sources and periods over which econometric studies have been carried out. But much of the explanation surely lies in the diversity of the countries themselves. Each has some special aspects of its history, geography, religion, society, polity, culture, and other features that influence how well or poorly a particular prescription for institutional reform or policy change will fare when applied to it.

Another way of saying this is that the econometric and theoretical studies are not the best way to generate policy prescriptions. Most cross-country regressions are a far from perfect fit; the myriad explanatory variables that have been tried explain only a fraction of the variance. Theoretical modeling explores the implications of one cause or mechanism in depth, deliberately isolating it from others, whereas policy prescriptions require that one consider all the different causes or mechanisms at work in a country and how they interact. The question policy prescribers must address, is not what creates success on average across countries, but what is going wrong in this country, and how can we put it right?

This seems to suggest a division of labor. Academic researchers should look for general tendencies and patterns and should not offer specific prescriptions for individual countries. That should be left to policy practitioners in the countries or in international organizations, who have much more detailed knowledge of the circumstances and needs of individual countries. Of course, some academics combine theoretical knowledge with deep knowledge of some countries; for example, Fafchamps (2004) examines many African countries through the lens of game-theoretic modeling of self-enforcing governance mechanisms and comes up with different prescriptions for three countries in different circumstances and at different stages of institutional evolution.7

How can practitioners combine their country knowledge with academic research? Must this be a totally ad hoc matter, or can one develop something of a system? I would like to suggest a framework building on the idea of “growth diagnostics” introduced by Hausmann, Rodrik, and Velasco (2005). Their basic idea is to identify the crucial or binding constraints in any situation and then to focus attention on remedying those. They develop a theoretical framework for this and operationalize it using a “tree” much like the branching decision diagrams in the trouble-shooting sections of product manuals. (Does the computer beep four times when you switch it on? If yes, skip to step 10; if no,...) Hausmann, Rodrik, and Velasco (2005, p. 2) ask practitioners to look at other symptoms that may be associated with low growth and use a succession of questions about symptoms to pin down a cause:

Is it inadequate returns to investment, inadequate private appropriability of the returns, or inadequate access to finance? If it is a case of low
returns, is that due to insufficient investment in complementary factors of production (such as human capital or infrastructure)? Or is it due to poor access to imported technologies? If it is a case of poor appropriability, is it due to high taxation, poor property rights and contract enforcement, labor-capital conflicts, or learning and coordination externalities? If it is a case of poor finance, are the problems with domestic financial markets or external ones?

This approach has much to recommend it. It looks at multiple dimensions of economic outcomes and tries to narrow down causes from a set of multiple possibilities. But its sequential mode of thinking based on a tree structure seems problematic. In reality, each case of development failure may have multiple causes acting simultaneously. A related but more general point concerns the idea of "diagnosis" itself. Dictionary definitions of the word are some variants of "identification of a disease or disorder based on a review of signs, symptoms, and laboratory findings." This sounds like a problem of Bayesian inference and is perhaps better captured in a table than in a tree.

Table 1 shows such a schematic representation. The first column has the various causes \( C_1, C_2, \ldots \) in its rows. Each cause could be a grouping or composite of several causes. And the list can be quite wide and comprehensive, including political and institutional obstacles as well as specific policy errors of commission or omission. The second column has the prior probabilities \( \pi_1, \pi_2, \ldots \) with which the various causes can occur. Historical analyses can provide estimates of these probabilities; specific knowledge of a country can help refine these estimates.\(^8\) The rest of the columns have as their headings various conceivable outcomes \( E_1, E_2, \ldots \); again, each outcome could be a cluster or composite of several outcomes. The entries in the cells are then the conditional probabilities, with \( P_{ij} \) being the probability that outcome \( E_i \) will occur when cause \( C_j \) is present. In principle, econometric research can yield such probabilities, although this is not how regression results are usually presented.

\[
\begin{array}{c|c|c|c|c}
\text{Causes} & \text{Prior probabilities} & \text{Effects} \\
\hline
C_1 & \pi_1 & E_1 & P_{11} & P_{12} \ldots & P_{1n} \\
C_2 & \pi_2 & E_2 & P_{21} & P_{22} \ldots & P_{2n} \\
& \vdots & \vdots & \vdots & \vdots \ddots & \vdots \\
C_m & \pi_m & E_m & P_{m1} & P_{m2} \ldots & P_{mn} \\
\end{array}
\]

\[^{8}\text{Avinash Dixit}\]
If we observe a particular effect, say \( E_7 \), then the Bayesian posterior probability that a particular cause, say \( C_5 \), is present becomes

\[
\frac{\pi_5 p_{5,7}}{\sum_{i=1}^{m} \pi_i p_{i,7}}
\]

If we want to be nearly certain whether a cause, say \( C_5 \), is present, we need to find an outcome, say \( E_7 \), which will more typically be a cluster of outcomes or symptoms and might be called a “syndrome,” such that

- It is very unlikely to occur when the underlying cause is any other cause, that is, \( p_{i,7} \) is close to zero when \( i \) is not equal to 5, so the posterior probability of \( C_5 \) is close to one conditional on observing \( E_7 \).
- It is very likely to occur when \( C_5 \) is present, that is, \( p_{5,7} \) is close to one, so the rest of the \( p_{5,i} \)'s are close to zero, and if some other effect is observed, the posterior probability of \( C_5 \) becomes close to zero.

The ultimate aim would be to find a complex of causes, or a syndrome, that satisfies both these conditions. Then we can reduce the table to a two-by-two case, where the syndrome we have identified is labeled \( C_1 \) and everything else is subsumed into \( C_2 \), whereas the outcome or effect we are trying to explain is \( E_1 \) and everything else is \( E_2 \). In that ideal setting, the conditional probabilities would be \( p_{11} = p_{22} = 1 \) and \( p_{12} = p_{21} = 0 \). Then the posterior probabilities are one for \( C_1 \) if \( E_1 \) is observed and also one for \( C_2 \) if \( E_2 \) is observed, so the outcome pins down the cause precisely. In reality, we are unlikely to get a situation where causes can be discerned from syndromes quite so well, but this provides an ideal to work toward.

I hope this very rough conceptual scheme can be developed into something more concrete and it proves useful, not only for thinking about the causes of some development problems, but also for identifying and then tackling them in practice. If this works, I will be delighted to have concluded a largely nihilistic tour of contradictions, conflicts, and confusion in the literature on a positive note.

Notes

Avinash Dixit is the John J. F. Sherrerd ’52 University Professor of Economics at Princeton University; his email address is dixitak@princeton.edu. This is a slightly revised version of a paper presented at the World Bank Development Economics Vice Presidency Lectures on April 21, 2005. It was intended to be provocative, not a balanced scholarly product. That spirit is preserved in the written version. For valuable comments and suggestions, the author thanks Philip Keefer (his discussant at the conference), Daron Acemoglu, Karla Hoff, Boris Pleskovic, and Dani Rodrik, who in the end probably disagree with much if not all of his interpretations of the literature.

1. However, others have rebutted this; for example, Mehlum, Moene, and Torvik (2006, p. 4) find that “the resource curse applies in countries with grabber friendly institutions but not in countries with producer friendly institutions.”

The World Bank Research Observer, vol. 22, no. 2 (Fall 2007)
2. However, Lamoreaux and Rosenthal (2005) argue that the institutional differences between France and the USA did not prevail during the nineteenth century, the crucial period for industrialization, but instead evolved slowly as a result of legislation during the 20th century.

3. Actually, Rodrik and Wacziarg find that either kind of transition is followed by better economic growth, but the effect is smaller for transitions from democracy to authoritarian rule. This suggests an alternative hypothesis. Suppose that purely exogenous random shocks are the main determinants of growth but that the public gives the rulers credit for good results and blames them for bad results. Rulers are overthrown when the results are particularly bad. Then the average improvement in growth following a regime change is just a reversion to the mean, and the effect is stronger for transitions to democracy because authoritarian rulers can be overthrown only when the economy performs abysmally whereas democratic governments may fall for smaller shortcomings.

4. One other feature of economic reforms in China is often highlighted and praised (for example, Qian 2003; Rodrik 2004b): China preserved established entitlements in an inframarginal manner, while liberalizing activity at the margin. For example, the state purchase system and price subsidies for agriculture were retained, thereby protecting government revenue and urban living costs, but gave farming households the freedom to sell the remaining output after meeting the state quotas at free market prices. However, this policy-decision is conceptually distinct from the issue of democracy or authoritarianism.

5. I restrict attention to institutions of property right protection and contract enforcement, but institutions and agencies for regulation of product and labor markets, taxation, trade, and so on, show an equally bewildering variety and equally complex considerations as to their relative merits. Islam (2004, p. 1) considers the institutions that affect international trade and concludes that “increasing trade integration is consistent with a wide array of institutional choices.” However, it is harder to specify which choice is best under what circumstances.

6. I have left out a few even more whimsical ideas, such as “The Mouse That Roared” recipe. In that movie, a small country in Europe, the Duchy of Grand Fenwick, declares war on the United States, with the idea that when Grand Fenwick loses the war, the United States will give it large amounts of aid. Unfortunately for the plan, Grand Fenwick wins.

7. Even such detailed and well thought-out combinations of academic and field research can be blown off course by unforeseen shocks. For example, Fafchamps regards Zimbabwe as having relatively good institutions and offers several suggestions for building on this base to improve financial intermediation. Just a few years later, Zimbabwe’s problems look quite different.

8. Medical diagnostics proceed using just such a combination of general statistical information about the relevant population and the doctor’s specific knowledge of an individual patient.

References


———. 2005. A Response to Albouy’s A Reexamination Based on Improved Settler Mortality Data. Cambridge, Mass.: Massachusetts Institute of Technology, Department of Economics.


Avinash Dixit


Comment on “Evaluating Recipes for Development Success”: The Policy Usefulness of Institutional and Political Analyses of Development

Philip Keefer

In “Evaluating Recipes for Development Success” Avinash Dixit criticizes recent efforts to identify the “fundamental” causes of development and to distill policy recommendations from these efforts. This comment focuses on the strand of that literature related to institutions and development. Two arguments are important: that the rule of law and the security of property rights are important for growth and that they are the product of political institutions. Professor Dixit argues that identification and other concerns undermine the second argument and inhibit the formulation of policy recommendations. While these concerns are valid, research has begun to disaggregate broad political institutions (democracy and autocracy) and to look at the details of political competition, such as voter information and politician credibility, which are both more robust determinants of political decision-making and more susceptible to policy interventions. JEL codes: O43, O17, O20, P30, P48.

Avinash Dixit reviews many of the recent contributions to the literature that examine the “big” questions in economic development, particularly those concerning the fundamental differences between countries that manage to sustain rapid economic growth and those that do not. In this compact and concise review Professor Dixit offers a critical assessment of this literature. He concludes by suggesting that academic development economists should refrain from offering country-specific policy prescriptions. Practitioners can nevertheless learn from the generalizations that academic research yields, but they should examine the plausibility of those generalizations, taking into account the many idiosyncrasies.
of the countries in which they work and the robustness of the evidence that supports the generalizations.

This comment focuses on those aspects of Professor Dixit's arguments that relate to the role of institutions in development. In this debate two questions are key. First, does the security of property rights (sometimes operationalized as the risk of expropriation, the enforceability of contracts with the government, or the extent of law and order) matter for growth? Second, what kinds of societies, institutions, or polities offer secure property rights? There is substantial uncertainty about answers to the second question. However, it is important to emphasize that there is substantially less uncertainty about the first. Moreover, answers to the second question currently emerging from the literature offer far more hope for policy relevance than is generally perceived.

### Does the Security of Property Rights Matter for Growth?

The key difficulty in testing theories linking the security of property rights (or anything else) to economic growth is the possibility that unobserved influences exist that simultaneously secure property rights and accelerate growth. Professor Dixit questions the validity of several efforts to solve this problem. These are based on particular historical and geographic features of countries that researchers theorize should determine the security of property rights but that should not directly affect growth. Professor Dixit argues that there are no robust econometric tests to validate these theoretical arguments. To the extent that each feature (settler mortality or European colonial heritage, for example) is used to test a corresponding theory linking historical factors to growth, this criticism is well-taken: the strength of the econometric validation of the particular theory depends on the power of the tests with which one judges whether the particular feature affects property rights but has no other effect on growth.

However, the strength of the empirical conclusion that the security of property rights is itself an important determinant of growth rests on a vast array of tests. It does not depend on the validity of particular theories on the genesis of secure property rights. The argument that the security of property rights determines growth has been tested with many instrumental variable strategies, with several different variables for measuring the security of property rights, over many different time periods, and with a huge array of control variables. Scholars have tested and rejected the possibility of reverse causality from growth to property rights. Micro evidence is also persuasive (for example, with respect to housing and agricultural investment). The relationship between property rights and
growth is, in other words, one of the most examined in the cross-country empirical development literature.

None of this means that the case for the security of property rights is iron-clad. It does mean that, by the standards of the Bayesian updating that Professor Dixit recommends that practitioners apply in evaluating conclusions from the academic literature, our priors are reasonably strong that the property rights–growth relationship is significant.

What Kinds of Institutions and Polities Offer Secure Property Rights?

What, though, are the practical implications of this relationship? What can policymakers do if their country exhibits insecure property rights—and what should donors do? Here, Professor Dixit identifies a widely shared frustration: even if he suspends his doubts about the econometrics, the academic literature offers no consensus on solutions, nor are many solutions susceptible to policy intervention. Countries cannot change their location or their history. Nevertheless, there is much greater consensus in the literature on one key point, either implicitly or explicitly: the security of property rights is a product of a country’s political environment. Which particular aspect of the political environment is a question at the forefront of current research. It is therefore not surprising that policy-relevant answers have been sparse.

Professor Dixit focuses on one aspect of the political foundation of secure property rights: the existence of competitive elections. Arguments in the literature that historical or geographic endowments influence property rights, to the extent that they are explicit about causal mechanisms, almost uniformly point to the influence of these endowments on the decisions of elites to allow competitive elections that empower nonelites (and thereby allow nonelites to protect their property from confiscation by elites). The democracy versus autocracy literature that Professor Dixit reviews is based on similar claims.

The argument linking democracy to property rights is straightforward. Elections, perhaps supplemented by political checks and balances, change the incentives of leaders by requiring them to seek the support of a majority of voters. Thus, leaders should be more inclined to provide higher quality public goods—not only secure property rights, but also better access to education and more efficient regulation. Moreover, elected leaders are more likely to be punished for expropriation, which slows growth and reduces the incomes of the majority. Unelected leaders, relying on a narrower support base to stay in power, have fewer incentives to provide public goods and are less likely to be punished for the slow growth engendered by expropriation.

Philip Keefer
Despite the clarity of the theory, the evidence linking democracy to the security of property rights (controlling for income per capita), or to growth more generally, is in fact quite mixed, as Professor Dixit points out. This reflects large variation in democratic performance. For example, secondary school enrollment was lower in more than 30 percent of democracies than in the median nondemocracy. More than 40 percent of democracies exhibit worse corruption than the median nondemocracy. And mean rule of law (through the 1990s) was almost the same in autocracies as in democracies.

The literature is fairly clear that competitive elections are an insufficient explanation for secure property rights. This is where Professor Dixit’s argument stops. However, political factors that might affect the security of property rights go far beyond competitive elections. The ways leaders attain and remain in office vary widely among both democracies and nondemocracies, and these variations have a significant effect on leaders’ incentives regarding public policy generally and the security of property rights specifically. For example, in some democracies leaders can make credible policy promises to a large fraction of the electorate; in others, to only a tiny fraction. In some democracies citizens are well-informed about what governments are doing; in others they are poorly informed. These political market imperfections disrupt mechanisms of electoral accountability. However, responding to the other important criticism that Professor Dixit raises, these other political factors are more amenable to change than, for example, history and regime type.

The variation in political incentives is easy to observe, even among democracies. In many democracies, members of the elected legislature spend all of their time on the private needs of constituents (placing them in government jobs, intervening on behalf of their children with educational authorities, and providing very local public goods). This job description is quite distant from that of a typical member of the German Bundestag, British Parliament, or U.S. Congress and reflects much greater pressure to provide narrowly targeted, private goods to key constituents and weaker pressure to provide high quality public goods benefiting all citizens. Analogous variation exists across autocracies. A cutting-edge area of research in the political economy of development—and in research on the determinants of secure property rights—is therefore precisely focused on why political incentives differ within regime type.

Recent work on democracies has examined four sources of variation: citizen information about government performance; the credibility of political competitors (the fact that the political party labels of candidates in Bangladesh, Ecuador, and Indonesia, for example, do not convey a credible reputation for broad policy preferences in the same way that political party identity does in mature democracies); social polarization, whether rooted in ethnic, income, or other differences; and variations in specific political and electoral institutions (first-past-the-post electoral rules provide different incentives to politicians than
proportional representation). Evidence links all of these to government incentives to provide public goods, including secure property rights (see Keefer 2004 and Keefer and Khemani 2005 for reviews). Just as important, compared with such determinants of political behavior as history and regime type, these sources of variation in political incentives have at least somewhat more tractable policy implications for what donors and governments should and should not do.

In this sense they have a direct bearing on Professor Dixit’s recommendation that policy measures should be incremental and sensitive to country characteristics. This recommendation is laudable. The question is, which characteristics? One inclination is to focus on the policy bottlenecks to development. However, reform almost always runs through government. Its approval and implementation therefore depend on political decision-making. Incremental approaches that fail to take the conditions of political decision-making into account in a systematic way are no more likely to succeed than “maximalist” approaches. The design and advocacy of practical and incremental reforms should therefore be informed by the political constraints on reform. Fortunately, the developing literature on political market imperfections, by moving beyond highly aggregated characteristics of countries, such as regime type, offers ways to think more precisely about political constraints in the context of reform strategy in specific country cases.

For example, in countries with significant credibility and information problems, reforms for which government officials can transparently take credit and that are easy for recipients to monitor are more likely to succeed. This situation favors efforts to improve the quantity of social services and may argue against large investments in quality, even where quality shortcomings are severe. On the other hand, in heavily polarized environments, where the precise distribution of benefits across groups is highly contentious, efforts to improve quality may be preferable. Thus, it may be preferable to pursue reforms for which the precise distribution of benefits is more ambiguous and potentially more even.

The analysis of political market imperfections also reveals where governments are so noncredible as to call into question their staying power and ability to resist overthrow, as in the case of fragile states. With fragile states there may be less incentive to launch elaborate programs of targeted assistance to the poorest members of the community. Elaborate programs do little for political credibility, when they obscure beneficiary perceptions of who is responsible for the benefits they receive—the political leadership of the country or lower level administrative agents. Less targeted programs, in which targeting is crude but easy to communicate and simple to implement, may offer a greater contribution to development by building political credibility, even at the cost of economic inefficiency.

Finally, this line of literature suggests political obstacles to reform that are more directly susceptible to reform than is regime type or history. The information contribution is particularly worthwhile. From the first Public Expenditure Tracking
Survey in Uganda, which led to a 90 percent reduction in the diversion of capitation grants to schools, to report cards on public services, pioneered in Bangalore, India, but expanding to China and elsewhere, a variety of tactics are emerging to close the information gap between citizens and politicians. Other strategies, such as the removal of restrictions on the media and increased transparency of government information, are also a direct response to information imperfections in political markets.

Professor Dixit's skepticism about the immediate policy usefulness of many of the "big ideas" in development is warranted, although incrementalism itself constitutes a big idea of sorts, one whose ultimate contribution to development remains unproven. Similarly, his insistence on country-based, analytically sound approaches to evaluating reform priorities is an essential lesson. But we should not lose sight of the "recipes" or big ideas. The analysis of political distortions in development decision-making emerges exactly from a concern for the fundamental drivers of development. Despite this—despite the fact that such analyses are concerned with big ideas—this line of research shows considerable promise in informing both the content and the design of the reform agenda in country-specific contexts.

Note

Philip Keefer is lead research economist in the Development Research Group at the World Bank; his email address is pkeefer@worldbank.org.

References


Domestic Bond Market Development: The Arirang Bond Experience in Korea

Jonathan A. Batten • Peter G. Szilagyi

This study contrasts the development of the Republic of Korea's market for won-denominated foreign bonds (Arirang) with similar markets in the Asia-Pacific region. It discusses the problems, concerns, and key issues related to the development of this market within the broader context of domestic, regional, and global bond market development. Korea's experience provides valuable lessons for other emerging market economies also seeking to build bond markets for local and foreign issuers. The sophistication of the local bond market is not enough to make it appealing to foreign borrowers. Market development demands ensuring an enabling infrastructure and a background of macroeconomic stability, nurturing local and international demand, deregulating capital flows, and minimizing exchange restrictions. JEL codes: F34, G18.

The 1997–98 East Asian financial crisis gave a significant boost to policy reform in the region, spanning the full spectrum of macroeconomic and microeconomic policy possibilities. Specific attention also focused on identifying why bank finance almost completely dominated financial markets in East Asia and on facilitating the development of local and regional bond markets. The idea was that more developed bond markets would make banking markets more efficient and competitive, and would help retain the region's vast household and corporate savings, which were directed largely to fixed-rate investment in Europe and the United States.¹ A two-tiered approach to financial market development aimed at both bank and bond market reform would also be complementary to longer term economic development, provided services could be delivered through efficient financial and legal institutions (Chakraborty and Ray 2006) and there was strong protection for investors and sound fiscal and monetary policy management by government (Burger and Warnock 2006b).²

These bond market reforms across East Asia have been quite successful. Bond market volumes have increased twofold or more, and corporate issuance has
expanded along with the government bond market. In addition, risk premiums, measured by spreads on sovereign eurobonds, have shrunk following a region-wide policy of structural reform matched by prudent fiscal management by governments. In several countries, including the Republic of Korea and Malaysia, the corporate bond market has caught up with that in the United States, as measured by a percentage of gross domestic product (GDP).

Despite these many improvements, however, the development of local financial markets remains modest by the standards of bond markets in developed countries, especially when viewed from a regional perspective. There is also considerable variation in the scale and scope of these markets that correspond on an anecdotal level with the legal jurisdictions in place: countries with common law-based legal systems seem to be more successful in developing their securities markets. Burger and Warnock (2006a) also find that emerging market economies with stronger legal institutions and better historical inflation performance have more developed local bond markets.

One critical market segment that remains largely overlooked in the region and that could help to elevate regional and domestic bond markets to the global plane advocated by McCauley and Park (2006) is the market for foreign bond issuers. These issuers include supranational organizations, such as the Asian Development Bank and the International Finance Corporation; prime-name corporations, such as IBM, Disney, and GE Finance; and banks and other financial institutions, such as Kreditanstalt für Wiederaufbau, that are well known for the diversity of their international bond offerings and that actively issue in many countries and currencies.

The only countries in Asia and the Pacific region that have consistently attracted these high-quality borrowers are the major financial centers of Singapore and Hong Kong, China, and the developed countries of Australia and Japan. Of the crisis economies that implemented radical regulatory change, only Korea has made significant progress in attracting bond issuance by nonresidents.

The objective of this article is to demonstrate the importance of the development of a domestic currency-denominated foreign bond market as the next stage in the bond market reform agenda. The article examines the case of Korea within the broader context of regional foreign bond market development. It builds on Hoschka (2005), who discusses the importance of multilateral development banks, especially the Asian Development Bank, in helping to expand nascent bond markets.

While each financial and bond market in the region differs to some extent, there are common problems that impede the development of corporate bond markets, including investor participation, liquidity, price transparency, credit ratings, and taxation (Leung 2006). One feature of Korea's experience that is especially valuable to other emerging market economies is that it was the first of
the crisis economies to actively pursue the expansion of its domestic bond market while also encouraging nonresident issuance through won-denominated "Arirang" bonds.

Today, the Arirang market is still modest in size, at less than 1.7 percent of corporate issuance, or around US$2.7 billion in 2006. This market developed without internationalization of the won or significant changes in capital and exchange controls, although recently further steps have been taken to internationalize the won. Reforms in these areas characterized the stellar development of the foreign bond market in Australia (McCauley 2006) and Singapore (Lian 2002). Nonetheless, the Arirang market has been integral to the development of nongovernment bond markets by encouraging the reform agenda, highlighting regulatory and infrastructure deficiencies, and helping to establish an investor base. Overall, Korea's experience holds important policy lessons for issuers, investors, regulators, and policymakers in other emerging market economies that are developing local bond markets.

The next section briefly discusses the expanding literature on financial market development with a focus on the role of bond markets. Then, the institutional context of Korea's financial markets is briefly described, with the focus on the corporate segment, where Arirang bonds belong from a regulatory perspective. Where possible, these market segments are also compared at a regional level. The problems, concerns, and key issues related specifically to the Arirang market are discussed next, followed by strategies for further development. The final section presents some lessons that may be applied to other financial markets.

The Need for Local Bond Markets

The rationale for bond market development appears clear. A sophisticated market reduces systemic risk and the probability of crisis, since the economy can then borrow in its own currency or in others but with longer maturities (Eichengreen and Hausmann 1999; Burger and Warnock 2006b). Better market mechanisms should aid risk sharing in the financial system and improve the ability to withstand prolonged shocks. The flow-on effects at the corporate level should lead to lower funding costs, improved resource allocation, more efficient corporate capital structures, and encouragement of innovation (Takagi 2002).

Much of the recent financial reform agenda in East Asia has focused on promoting corporate bond market development to expand the financing options available to the corporate sector (Herring and Chatusripitak 2000). McCauley and Park (2006) note three aspects of this vision of well-developed corporate bond markets: a series of domestic markets in which domestic investors provide funds to domestic issuers, a regional bond market denominated in regional...
currencies with regional investors and issuers, and a global market in which East Asian borrowers and possibly investors are minor players. McCauley and Park argue that the ultimate objective of the development of national bond markets should be integration into a global market. Development of a foreign bond market is consistent with such a vision.

Initially, attention was paid to development of domestic bond markets, both corporate and government (Batten and Kim 2001). Many early government reforms focused on the need to build infrastructure, including settlement systems (Park and Rhee 2006), and to establish reputable credit ratings (Kisselev and Packer 2006) and benchmark yield curves (Woodbridge 2001).

Later, attention shifted to regional markets. The Changmai Proposal at the meetings of the Asian Cooperation Dialogue in June 2003 called for the development of an Asian regional bond market (Pei 2005) and an Asian Bond Fund (Leung 2006). Park and Park (2004) advocated a market-led approach to bond market development coupled with domestic financial reform, to allow the development of viable domestic bond markets before attempts to tackle regional bond market development. Local derivatives markets would also be needed to facilitate risk and maturity transformation (Burger and Warnock 2006b). Only recently has attention turned to the importance of foreign participation in domestic markets (Burger and Warnock 2006a), which includes the issuance activities of multilateral development banks (Hoschka 2005) and foreign investors (Bae, Yun, and Bailey 2006).

Bonds issued in regional currencies, or against a basket of local currencies as suggested by Ito (2004), minimize the double mismatch problem (exchange and maturity mismatch between assets and liabilities) that local bond issuers usually experience. Historically, local issuers tend to issue in the major currencies (U.S. dollars, yen, and euro), and then either swap the proceeds into local currency (interest rate parity theory suggests this should deliver funds equivalent in yield to what is available in the domestic market) or, more often, sell the foreign currency proceeds in spot foreign exchange markets, leaving the repayment cash flows unhedged. This strategy is consistent with a carry-trade, which relies on the unhedged funding of high-yielding assets in foreign currency with low coupon currencies such as the yen to deliver speculative profits—provided the high-yielding currency does not depreciate below the interest rate differential between the two. The risk of these unhedged borrowings is that depreciation of the local currency can destroy the equity position of the local borrowing firms, as happened throughout the region (especially in Indonesia) during the East Asian financial crisis.7

Possibly more important for the post-crisis economies of East Asia, competition for borrowers reduces the dependence of firms on banking relationships (Weinstein and Yafeh 1998) and may induce banks to lend to lower quality borrowers than
otherwise (Dinc 2000). Competitive pressure from the bond market encourages banks to monitor their credit decisions more effectively when information is asymmetric (Stulz 2000) or when there is poor investor protection (Modigliani and Perotti 2000). There is now consensus that banks and markets can coexist efficiently even in bank-oriented financial systems (Levine 1997; Boyd and Smith 1998; Bolton and Freixas 2000; Ongena and Smith 2000; Allen and Santomero 2001). Chakraborty and Ray (2006) recently established that although stronger bank monitoring helps to resolve information asymmetries and agency concerns, it is the efficiency of financial and legal institutions that influences growth outcomes, whether there is a bank- or a market-based financial system. This finding is consistent with Burger and Warnock (2006b), who note the importance of the legal setting, especially creditor rights, for bond market development.

Empirical evidence suggests that, in East Asia at least, bond markets should be larger and more developed than they are (Eichengreen and Luengnaruemitchai 2004; Lejot, Arner, and Liu 2006). The list of obstacles to be overcome is extensive, as a range of studies have pointed out (Benzie 1992; Emery 1997; Schinasi and Smith 1998; Kim 1999; Batten and Kim 2001; IMF 2005; Lejot, Arner, and Liu 2006; Leung 2006). Among them are the need for enabling regulation, including reform of withholding and other foreign investor taxes (Lejot, Arner, and Liu 2006); continuing reform of corporate governance, which includes better creditor rights, bankruptcy procedures, and contract enforcement (Beck, Levine, and Loayza 2000; Burger and Warnock 2006); and strong financial infrastructure for better information disclosure, the establishment of reliable credit ratings (Kisselev and Packer 2006) and robust benchmark yield curves; and high-quality settlement and risk management systems (Rhee 2003).

The institutional setting before the East Asian financial crisis clearly favored unhedged foreign borrowing with short-dated maturities over domestic-sourced and higher yielding bank debt or securities. Overvalued currencies in protected exchange rate regimes contributed to the bias. There is also some evidence of agency effects from numerous family-owned corporations, which, to preserve information asymmetries, sought bank financing instead of securities issued in either domestic or international markets.

The development of foreign bond markets is important for several reasons. Initially, it serves as a barometer of general development in the local bond market through the availability of better quality and longer dated securities that offer improved diversification for local investors (Jiang and McCauley, 2004). Foreign issuers may introduce best practice in issuance, disclosure, and documentation. And with supranationals such as the Asian Development Bank and the International Finance Corporation, which are often the first foreign bond issuers in developing markets, there is often the prospect of altruistic goals of enhancing bond market development more generally.
The Korean Bond Market

The foundations of the bond market for Korean government securities were laid in the early 1950s. Corporate bonds appeared in 1963, but issues were restricted to short maturities and were effectively dependent on bank guarantees. Not until the mid-1990s did the market begin to open to foreign investors, and not until 1997 was it fully liberalized, spurred by the currency crisis (Noland 2005). From the late 1980s growth of the government bond market had been slowed by continual current account surpluses. Thus, instead of a robust treasury bill market, quasi-government securities were issued to assist with monetary and exchange rate stability and housing development. At this time chaebol-issued corporate bonds constituted by far the largest segment of the Korean bond market.

The currency crisis brought about fundamental changes in the market. The government had to raise huge amounts of funds for fiscal stimulation and financial restructuring, which induced rapid expansion in all segments of the public bond market. In 1998, treasury issues alone increased sixfold to won 12.5 trillion. Corporate issues also jumped to a staggering won 56 trillion. Firms had to shift borrowing to nonguaranteed securities, as troubled financial institutions were reluctant to extend credit lines or provide credit guarantees. Large quantities of asset-backed securities were issued simultaneously to securitize nonperforming loans and credit card receivables, creating one of the most sophisticated structured finance markets in the region. At the same time the huge surge of fund inflows into investment trust companies secured ample demand for these securities.

The infrastructure for the market was built gradually. Market operations are overseen by the Ministry of Finance and Economy and the Financial Supervisory Commission. Since 1998 the Bank of Korea has had only indirect oversight through supervision of payment and settlement systems and foreign exchange reserves. The securities market is largely self-regulated through organizations such as the Korea Securities Dealers Association, the Korea Exchange, and the Korea Securities Depository, and four local agencies assign credit ratings: Korea Investor Service (a Moody's affiliate), Korea Ratings (a Fitch affiliate), National Information & Credit Evaluation, and Seoul Credit Rating & Information. The underwriting market has also grown competitive, with Dealogic Data showing 10 bookrunners with a market share of at least 3.5 percent, the two largest being the Korea Development Bank and Woori Finance, a subsidiary of Woori Bank.

Today, Korea’s bond market is the second largest in East Asia (table 1). In June 2005 it was valued at US$599.8 billion, or 81 percent of GDP, a nearly fivefold increase over 1997, an astonishing rate of growth. It is also the most diverse in the region, with corporate securities accounting for 26 percent of its volume and financial institution bonds for 44 percent. Nontreasury public bonds and asset-backed securities add another US$400 billion to the value of the market.
Quasi-government securities include monetary stabilization bonds, foreign exchange stabilization bonds, national housing bonds, and special public bonds issued largely to finance infrastructure improvements in transport, electricity, and telecommunications.

The market is unconcentrated but only moderately diverse, as issuance is dominated by chaebols and their subsidiaries. Accordingly, credit quality in the market is very high and increasing in local terms (table 2). Bonds tend to be straight, unsecured, and almost exclusively nonguaranteed. Equity-linked instruments, predominantly convertibles, are relatively rare, with only 42 listed on the Korea Exchange. Of the more than 2,400 corporate bonds listed on the Korea Exchange, more than 90 percent are unsecured straight issues. Most securities are listed, largely because of restrictions on institutional investors investing in unlisted bonds. Corporate issues are concentrated at the shorter end of the maturity spectrum, with three-year bonds the most popular.

The low concentration of issuance limits liquidity to some extent. In 2005 the average issue size was about US$40 million. Turnover in secondary markets is still relatively high at 3.3 times the outstanding amount in government bonds and 1.0 time the outstanding amount in corporate securities the highest in East Asia. Trading in corporate bonds has declined by nearly 75 percent since 1999 (table 3). As in most countries, most corporate bond trades take place largely over the counter, administered mainly through the Korea Securities Dealers Association Free Board. Only 1.6 percent of trades are accounted for by the Korea Exchange, which reflects the low standardizability of corporate issues.

Jonathan A. Batten and Peter G. Szilagyi
Table 2. Outstanding Bonds in Korea, by Type (billions of won)

<table>
<thead>
<tr>
<th>Bond type</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>547,449</td>
<td>600,139</td>
<td>650,110</td>
<td>736,160</td>
<td>805,776</td>
</tr>
<tr>
<td>Government</td>
<td>78,126</td>
<td>95,808</td>
<td>137,677</td>
<td>197,932</td>
<td>234,126</td>
</tr>
<tr>
<td>Treasuries</td>
<td>46,032</td>
<td>52,389</td>
<td>82,427</td>
<td>142,421</td>
<td>180,855</td>
</tr>
<tr>
<td>Foreign exchange stabilization bonds</td>
<td>8,700</td>
<td>15,850</td>
<td>23,650</td>
<td>22,200</td>
<td>15,300</td>
</tr>
<tr>
<td>National housing bonds</td>
<td>21,268</td>
<td>26,469</td>
<td>31,001</td>
<td>33,311</td>
<td>37,376</td>
</tr>
<tr>
<td>Special public</td>
<td>134,343</td>
<td>133,417</td>
<td>118,286</td>
<td>115,288</td>
<td>117,191</td>
</tr>
<tr>
<td>Monetary stabilization bonds</td>
<td>79,121</td>
<td>83,890</td>
<td>105,497</td>
<td>142,729</td>
<td>155,235</td>
</tr>
<tr>
<td>Financial debentures</td>
<td>83,660</td>
<td>120,898</td>
<td>123,963</td>
<td>134,854</td>
<td>145,523</td>
</tr>
<tr>
<td>Municipal</td>
<td>9,246</td>
<td>8,954</td>
<td>9,919</td>
<td>10,553</td>
<td>11,210</td>
</tr>
<tr>
<td>Corporate</td>
<td>162,953</td>
<td>157,172</td>
<td>154,748</td>
<td>134,804</td>
<td>142,492</td>
</tr>
<tr>
<td>Guaranteed</td>
<td>5,987</td>
<td>4,862</td>
<td>3,364</td>
<td>755</td>
<td>323</td>
</tr>
<tr>
<td>Secured</td>
<td>64,060</td>
<td>63,454</td>
<td>63,121</td>
<td>48,878</td>
<td>40,704</td>
</tr>
<tr>
<td>AAA</td>
<td>9,384</td>
<td>15,618</td>
<td>16,369</td>
<td>15,542</td>
<td>14,583</td>
</tr>
<tr>
<td>AA</td>
<td>16,862</td>
<td>16,219</td>
<td>18,259</td>
<td>23,604</td>
<td>31,633</td>
</tr>
<tr>
<td>A</td>
<td>19,232</td>
<td>20,145</td>
<td>21,431</td>
<td>22,851</td>
<td>29,664</td>
</tr>
<tr>
<td>BBB</td>
<td>20,565</td>
<td>16,823</td>
<td>15,779</td>
<td>16,008</td>
<td>17,686</td>
</tr>
<tr>
<td>BB</td>
<td>7,855</td>
<td>6,446</td>
<td>5,506</td>
<td>1,251</td>
<td>1,075</td>
</tr>
<tr>
<td>B</td>
<td>4,454</td>
<td>3,589</td>
<td>1,592</td>
<td>457</td>
<td>432</td>
</tr>
<tr>
<td>CCC-D</td>
<td>3,545</td>
<td>1,178</td>
<td>536</td>
<td>345</td>
<td>291</td>
</tr>
</tbody>
</table>

Source: Korea BondWeb.

The remarkable decline in corporate bond trading is deceptive, however, as it simply reflects turnover settling down after a series of runs between 1999 and 2003. Ferocious trading in the market was fed by three major shocks. Each of

Table 3. Government and Corporate Bond Trading in Korea (billions of won)

<table>
<thead>
<tr>
<th>Year</th>
<th>Government bonds</th>
<th></th>
<th></th>
<th>Corporate bonds</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Over the counter</td>
<td>Exchange</td>
<td>Exchange</td>
<td>Over the counter</td>
<td>Exchange</td>
<td>Exchange</td>
</tr>
<tr>
<td></td>
<td>counter</td>
<td>trading</td>
<td>trading</td>
<td>counter</td>
<td>trading</td>
<td>trading</td>
</tr>
<tr>
<td></td>
<td></td>
<td>share (%)</td>
<td>share (%)</td>
<td></td>
<td>share (%)</td>
<td>share (%)</td>
</tr>
<tr>
<td>1997</td>
<td>15,848.6</td>
<td>237.4</td>
<td>1.5</td>
<td>143,423.1</td>
<td>3,807.1</td>
<td>2.6</td>
</tr>
<tr>
<td>1998</td>
<td>44,554.8</td>
<td>6,519.9</td>
<td>12.8</td>
<td>391,951.9</td>
<td>8,968.9</td>
<td>2.2</td>
</tr>
<tr>
<td>1999</td>
<td>397,504.4</td>
<td>281,921.7</td>
<td>41.5</td>
<td>442,891.0</td>
<td>11,685.0</td>
<td>2.6</td>
</tr>
<tr>
<td>2000</td>
<td>582,662.0</td>
<td>23,521.2</td>
<td>3.9</td>
<td>272,401.0</td>
<td>3,648.4</td>
<td>1.3</td>
</tr>
<tr>
<td>2001</td>
<td>985,028.0</td>
<td>12,213.6</td>
<td>1.2</td>
<td>263,376.0</td>
<td>2,012.6</td>
<td>0.8</td>
</tr>
<tr>
<td>2002</td>
<td>731,811.0</td>
<td>46,062.8</td>
<td>5.9</td>
<td>223,808.0</td>
<td>1,111.4</td>
<td>0.5</td>
</tr>
<tr>
<td>2003</td>
<td>1,001,362.8</td>
<td>214,009.9</td>
<td>17.6</td>
<td>164,858.8</td>
<td>892.0</td>
<td>0.5</td>
</tr>
<tr>
<td>2004</td>
<td>1,524,504.0</td>
<td>383,122.5</td>
<td>20.1</td>
<td>140,560.0</td>
<td>986.7</td>
<td>0.7</td>
</tr>
<tr>
<td>2005</td>
<td>1,550,280.0</td>
<td>360,824.4</td>
<td>18.9</td>
<td>115,752.0</td>
<td>1,935.1</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Source: Korea BondWeb.
these was triggered by a sudden downgrade (Daewoo, Hyundai, SK Group, and LG Card), followed by a run on investment trust companies by households and firms, distress sales of bonds, and eventually government intervention (Jiang and McCauley 2004). Corporate issuance has never recovered, as firms could no longer delay much needed deleveraging. This encourages hope that future growth in the market will be underpinned by much healthier conditions.

Vast changes have occurred in the corporate bond investor base since 1998 (table 4). In 1998 the huge rush of capital inflows helped investment trust companies lift their market share to 62 percent, but by September 2005 the reversal of capital flows reduced their share to just 14 percent. Primarily yield-driven investors, such as small mutual savings banks and credit unions, reflecting their appetite for high returns in exchange for modest default risk, absorbed the disposals. The corporate bond holdings of banks and insurance firms have also increased over time.

On the whole, market conditions in Korea remain unattractive to foreign investors, who hold less than 0.5 percent of both corporate bonds and treasuries. This is only partly due to the market’s vulnerability to volatility spillover effects from regional financial markets and past capital controls. Because of withholding tax considerations, foreign participation occurs in the forward market rather than the spot market. More important, global portfolio managers stay away because the Korean won is not internationalized and offers no diversification benefits. As a result, even the government’s ability to borrow abroad in won is severely limited—Eichengreen, Hausmann, and Panizza (2003) call this the “original sin.”

Table 4. Corporate Bond Investors in Korea (percent)

<table>
<thead>
<tr>
<th>Investor</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banks</td>
<td>5.3</td>
<td>4.9</td>
<td>8.1</td>
<td>12.8</td>
<td>12.8</td>
<td>14.3</td>
<td>12.2</td>
<td>11.8</td>
</tr>
<tr>
<td>Bank trusts</td>
<td>10.3</td>
<td>10.4</td>
<td>7.4</td>
<td>6.4</td>
<td>8.1</td>
<td>6.3</td>
<td>5.5</td>
<td>5.4</td>
</tr>
<tr>
<td>Investment trust companies</td>
<td>62.4</td>
<td>60.2</td>
<td>29.5</td>
<td>20.8</td>
<td>18.7</td>
<td>11.9</td>
<td>15.3</td>
<td>13.9</td>
</tr>
<tr>
<td>Security companies</td>
<td>8.7</td>
<td>12.3</td>
<td>15</td>
<td>13.8</td>
<td>15.5</td>
<td>20</td>
<td>21</td>
<td>22.8</td>
</tr>
<tr>
<td>Insurance companies</td>
<td>2.1</td>
<td>1.5</td>
<td>5.5</td>
<td>7.2</td>
<td>8.9</td>
<td>11.8</td>
<td>11.8</td>
<td>11.8</td>
</tr>
<tr>
<td>Pension funds</td>
<td>0</td>
<td>5.6</td>
<td>9.2</td>
<td>10.7</td>
<td>10.4</td>
<td>9.1</td>
<td>6</td>
<td>5.2</td>
</tr>
<tr>
<td>Mutual savings banks, credit</td>
<td>11.2</td>
<td>5</td>
<td>25.3</td>
<td>28.4</td>
<td>25.6</td>
<td>26.5</td>
<td>28</td>
<td>29.1</td>
</tr>
<tr>
<td>unions, individuals, others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

General Conditions in the Arirang Market

Arirang bonds could be an important vehicle for promoting the internationalization of Korea's capital markets. A key attraction of foreign-issued bonds is their ability to assist the currency matching of Korean assets, which should significantly encourage cross-border investment in the medium term. The potential of the market is further amplified as a channel for the vast savings accumulated in Korea's financial system. Credit-worthy foreign institutions could raise cheap Korean funds through local won-denominated issuance and then swap these, through cross-currency and interest rate swaps, into the currency of their choice when windows of opportunity appear in domestic and international markets.

Currently, the Arirang market remains unable to serve the financing needs of foreign borrowers and constitutes barely 0.8 percent of the corporate bond market. Korean bonds and the won have very little appeal to foreign investors, although efforts to internationalize the won are being accelerated. Eichengreen (2004) points out that this may be a result of Korean monetary policy, which has restricted the flow of capital and intervened extensively to support the U.S. dollar. In addition, withholding taxes of 25 percent on foreign investors' bond interest income compares unfavorably with taxes in important regional markets such as Hong Kong, China, and Singapore. The Arirang market faces no competition from won eurobonds, as Japan's Samurai market does, and nonresident issuance directly in a foreign currency is not yet a viable option (Kimchi bonds, denominated in U.S. dollars, are still in the development stage).

The first foreign borrower to tap Korea's domestic market was the Asian Development Bank in 1995, with a seven-year issue serviced by the Korea Development Bank (see Hoschka 2005 for further information on multilateral development bank financing). Two more supranational organizations followed in 1997, but they have been absent from the market since then. Not until 1999 were foreign firms allowed to issue Arirang bonds, pursuant to the Foreign Exchange Market Liberalization Act of the same year. In 2006 there were six foreign bonds denominated in euros and U.S. dollars, totaling US$580 million. The two U.S. dollar issues were by the U.S.-based Bear Stearns Co and were floating-rate issues, while the four euro issues were both fixed- and floating-rate issues. (Details on all Arirang bonds issued as of year-end 2006 are provided in appendix table A1).

Comparison of the Korea Foreign Bond Market with Others in the Region

Of the four markets investigated in the Asia and Pacific region—Australia; Hong Kong, China; Singapore; and Korea—Australia has the largest foreign bond market with a total issuance of approximately US$70.7 billion (A$90.7) through
236 issues (table 5). The foreign bond markets in Hong Kong, China (US$2.4 billion) and Korea (US$2.7 billion) are approximately the same size and have had a similar number of issues (52 and 55). Despite the sophistication of the financial markets in Singapore, the foreign bond market is comparatively small in value (US$1.3 billion) and in the number of issues (19). Singapore sought to address

### Table 5. Foreign Bond Issues, Selected Asia and Pacific Economies June 1990–December 2006

<table>
<thead>
<tr>
<th>Foreign bond characteristics</th>
<th>Number</th>
<th>Share of total (%)</th>
<th>Number</th>
<th>Share of total (%)</th>
<th>Number</th>
<th>Share of total (%)</th>
<th>Number</th>
<th>Share of total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating (Standard &amp; Poor's and Moody's)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>76</td>
<td>32</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>AA</td>
<td>73</td>
<td>31</td>
<td>12</td>
<td>23</td>
<td>9</td>
<td>47</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>A</td>
<td>42</td>
<td>18</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>10</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>BBB</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>24</td>
<td>44</td>
</tr>
<tr>
<td>BB</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not rated/Not available</td>
<td>42</td>
<td>17</td>
<td>34</td>
<td>65</td>
<td>7</td>
<td>38</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>Coupon type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed</td>
<td>145</td>
<td>62</td>
<td>28</td>
<td>54</td>
<td>17</td>
<td>90</td>
<td>50</td>
<td>91</td>
</tr>
<tr>
<td>Floating</td>
<td>78</td>
<td>33</td>
<td>20</td>
<td>38</td>
<td>2</td>
<td>10</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selling features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private placement</td>
<td>3</td>
<td>1</td>
<td>25</td>
<td>48</td>
<td>0</td>
<td>7</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Exchange listed</td>
<td>97</td>
<td>41</td>
<td>12</td>
<td>63</td>
<td>12</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other/not available</td>
<td>135</td>
<td>55</td>
<td>27</td>
<td>52</td>
<td>7</td>
<td>37</td>
<td>36</td>
<td>67</td>
</tr>
<tr>
<td>Currency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local</td>
<td>235</td>
<td>99</td>
<td>42</td>
<td>81</td>
<td>18</td>
<td>95</td>
<td>49</td>
<td>89</td>
</tr>
<tr>
<td>Foreign currency</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>19</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Total bonds issued</td>
<td>236</td>
<td>100</td>
<td>52</td>
<td>100</td>
<td>19</td>
<td>100</td>
<td>55</td>
<td>100</td>
</tr>
<tr>
<td>Amount issued, local currency (billion)</td>
<td>A$90.7</td>
<td>HK$19.05</td>
<td>US$2.0</td>
<td>W486</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average issue size (million)</td>
<td>A$384</td>
<td>HK$366</td>
<td>S$105</td>
<td>W51,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Reuters Fixed Income Database.
the absence of a critical mass of corporations with the size and credit quality to issue domestic bonds by encouraging foreign multinationals and multilateral institutions to issue local currency bonds. Of the 19 foreign bond issues recorded, 17 are by financial institutions.

Foreign bonds in nondomestic currency (usually U.S. dollars and euros) remain a small submarket, at best described as experimental in the case of Australia and Singapore, although a modest number of issues have now been undertaken in Korea and Hong Kong. In Australia international investors are likely to maintain unhedged Australian dollar currency positions since they have the buffer of the positive interest rate premium (McCaulley 2006). Speculators may also be funding these higher yielding asset positions through low-yielding carry-trades (for example, unhedged borrowings in Japanese yen).

The Australian market is a high credit-quality market (AAA, AA, and A-rated account for 82 percent of issues), dominated by large bond tranches of about US$300 million. These bonds often substitute for risk-free government securities in the portfolios of pension funds in Australia. The average size of the bond issues in the other three markets is less than US$70 million, and while credit quality is also high in the Hong Kong and Singapore markets, Korea is an exception, with many bonds issued at just investment grade (24 BBB, or 44 percent). A significant proportion of bonds are listed on local exchanges in Singapore (63 percent) and Australia (41 percent). In Hong Kong many issues (25, or 48 percent) are privately placed and so do not require credit ratings. Private placements do appear to be an important selling feature in the other markets. Fixed-rate coupons dominate all markets, with the greatest concentration in Korea (91 percent) and Singapore (90 percent); Australia (62 percent) and Hong Kong (54 percent) are more balanced with respect to floating-rate issues. Few bonds are issued with option features (calls, puts, or convertibles) in any market, although five Arirang bonds issued by the U.S. subsidiary of SKC had put (option) features.

While there is no evidence of any systematic or institutional discrimination against foreign bond issuers in the Arirang market, issuance has been dominated by the foreign subsidiaries of Korean chaebols, which tap the already familiar market to achieve favorable funding conditions. These institutions are well known to local institutional investors even though they may not be the most creditworthy. This is in sharp contrast with other foreign bond markets in the region, where supranationals and genuine nonresident borrowers of typically high credit quality account for the bulk of issuance.

**Potential Risks from Expansion of the Arirang Market**

Before discussing the roots of the underdevelopment of the Arirang market, it is important to deal briefly with some of the concerns about the negative
implications of the expansion of the Arirang market, which are largely unfounded. For example, there is no evidence that foreign bonds would crowd out local offerings. In fact, the reverse may be the case. In the yield-driven Korean market the significant sector issues by local firms and financial institutions may inhibit price opportunities for foreign borrowers, which would typically be on the quality end of the market. This problem could be mitigated in a number of ways. For example, coordinated regional central bank investment in foreign-issued securities could stimulate demand and absorb much of the U.S. dollar-denominated reserve buildup in the region. A framework similar to the Asia–Europe Meeting Trust Fund set up to assist East Asia’s post-crisis economies would probably serve this purpose well.

Another potential problem relates to Korea’s relatively thin foreign exchange market. Foreign exchange trading in Korea is increasing, but it remains modest compared with that of other similar-size economies. Compare daily spot turnover of US$19.8 billion in Korea with US$48.9 billion in the Netherlands, US$80.8 billion in Australia, US$102 billion in Hong Kong, and US$125.3 billion in Singapore. Over-the-counter derivatives turnover of US$10.3 billion a day is also modest (BIS 2005). Evidence from other financial markets suggests that most foreign bond issues are swapped into foreign currencies. The same holds in Korea, where chaebol subsidiaries tap the local market and then swap the proceeds into U.S. dollars. If these transactions are relatively large, they could enhance exchange rate volatility and exert downward pressure on the won.

A simple econometric analysis (an event study) of the impact that Arirang issues have had on the dollar–won exchange rate in the period up to June 2006 can help to show whether this is a significant issue. Over 0- to 15-day event windows Arirang issuance has no clear effect in terms of cumulative returns or excess volatility (figures 1a and 1b). For example, a comparison of volatility and returns five days before and five days after issuance shows no statistically significant differences. Of course, the sample and issues are both small, but so too is the market, and the confidence interval at the 95 percent level includes most observations. When cumulative returns are plotted, there does appear to be a small spike in positive returns on day 0 (+0.0005) which persists although a 95 percent confidence interval includes all points five days before and after (figure 1c). To the extent that price discovery in markets is usually associated with a spike in volatility around the event date, this simple analysis suggests no statistically significant volatility spikes when Arirang bonds are issued.

Repeating this experiment for other markets is difficult because of the number of issues and their interaction with other external factors. In Singapore, institutions have historically been reluctant to short the local currency because of reporting requirements and meticulous oversight by the monetary authority. Thus, it is unlikely that the liquidity effects of foreign bond issues would be
Figure 1. Does the Issuance of Arirang Bonds Destabilize the U.S. Dollar/Won Exchange Rate?
(A) Average standard deviation of returns around 0- to 15-day event windows. (B) Average returns (log of daily change) around 0- to 15-day event windows. (C) Average cumulative returns (log of daily change) around 0- to 15-day event windows. Source: Authors' analysis based on data from Reuters' Fixed Income Database.
arbitraged. In Hong Kong, this is an irrelevant issue given the peg to the U.S. dollar. In Australia and Japan, with large offsetting capital and trade flows and well-developed and informed foreign exchange markets, any such effect would be hard to imagine.

**Economic and Financial Conditions Surrounding the Arirang Market**

Development of the Arirang market is obviously dependent on expectations about Korea’s economy as a whole. Burger and Warnock (2006a) emphasize the importance of a climate of macroeconomic stability to facilitate local bond market development. Domestic economic growth is expected to maintain a moderate rate of growth of 4–5 percent, which should underpin demand and supply in both the corporate and treasury segments of the bond market. The overall picture of treasury issuance is an interplay between proposed fiscal initiatives that would add to the expanding government deficit and concerns over the current account surplus and the buildup of foreign reserves. Korea’s foreign reserves are the fifth largest in the world, at US$240 billion in January 2007, and continue to be fed by current account surpluses, estimated at around US$17 billion or 2 percent of GDP.

There are signs of improved confidence by foreign investors. Inbound foreign direct investment (FDI) was US$11.6 billion in 2005, and Korea has broken into the top 20 in FDI attractiveness on the list compiled by the United Nations Conference on Trade and Development. High value-added FDI dominates, and most inbound investment is through mergers and acquisitions. Coupled with foreign portfolio investment, foreign shareholdings have risen to 45 percent of market capitalization on the Korea Exchange. Capital inflows are only partially offset by outbound FDI, which the government encourages to reduce upward pressure on the won. The government is also using foreign exchange stabilization bonds to prevent the won from appreciating, but the won now trades consistently below 1,000 to the U.S. dollar for the first time since 1997.

To date, there is little evidence that foreign firms would finance FDI-related expenditure through Arirang issuance, but expected future inflows of FDI could provide a basis for foreign participation in the market. First, however, the proper enabling environment needs to be in place. The environmental conditions that are critical to the development of a robust Arirang market are discussed below.

**Maintaining Sustainable Economic Growth and Reducing Default Risk.** A stable economic situation and reduced corporate leverage create expectations of earnings upgrades among corporate borrowers. The reduction in corporate debt to equity ratio has been remarkable, dropping from 182 percent in 2001 to 104 percent in 2004, and the default ratio has fallen below 0.1 (Lee and Kim 2006).
Since 2004 credit spreads have remained near historical lows, and in the BBB segment they have narrowed spectacularly. This makes the market more attractive to both domestic and foreign issuers. However, it is likely to reduce yield-oriented investor demand for Arirang issuance by top quality international borrowers.

Accommodating Declining Savings and High Levels of Consumer Credit. Korea's savings rate has stabilized in recent years, but it is on the decline from a long-term perspective. It dropped dramatically between 1998 and 2002, from 38 percent to 30 percent, due to credit card abuse, which triggered the collapse of LG Credit Card in 2003. To some extent, this means that household capacity to invest in securities other than through reinvestment or asset substitution remains limited. A similar story emerges from the low level of quasi money in Korea—at 62 percent of GDP in 2004, well below levels in many other countries in the region, including China, Malaysia, and Thailand. The supply of quasi money, which comprises currency not deposited in bank accounts and demand deposits of the central bank, continues to grow, however, and real income growth has been high in Korea, behind only China and Vietnam since 1998.

Improving Corporate Governance. Korean efforts to improve corporate governance have yielded some success, but they have also been heavily criticized. Korea's governance practices are ranked sixth among 10 Asian countries by the Asian Corporate Governance Association and CLSA Asia-Pacific Markets (ACGA and CLSA, 2005). Compliance with international accounting and auditing standards is high, and Korea was the first country in the region to pass a law allowing class action lawsuits for securities cases. However, it ranks less well on regulations and their implementation, the quality of regulatory and market-based enforcement, and the political and regulatory environment. Korea also scores lower on the World Bank governance indicators than do many of its regional counterparts, including Japan, Taiwan, China, Singapore, and Hong Kong, with improvements called for in particular in regulatory efficiency, rule of law, and control of corruption.11

Reforms in the governance culture of Korean firms are also criticized as largely superficial. Ownership structures are opaque, and independent institutional investors remain unorganized. Standard & Poor's (2006) has indicated that complex family-centered ownership and support to noncore subsidiaries harm corporate credit ratings. The efforts of governance-aware professionals, academics, and activists such as the People’s Solidarity for Participatory Democracy have brought increasing attention to these issues.

Maintaining Stability in Monetary Policy and the Exchange Rate. The domestic, economic, and financial environments are favorable overall to potential issuers in the
Arirang market. Price stability is good, with domestic demand pressures remaining low as households are deleveraging and credit delinquency is falling. More concern is voiced over the recent appreciation of the won. Unhedged foreign borrowers would be reluctant to tap Korea's domestic market if the won were likely to appreciate beyond the interest rate differential with the issuer's reporting currency. In 2006 the won traded between 1,150 and 1,240 against the euro and 990 and 910 against the U.S. dollar, gaining some 4.3 percent against the dollar year on year. The dollar exchange rate is nonetheless stable, with volatility at its lowest in the post-crisis period. Further rises in the exchange rate are also expected to remain modest, as indicated by the small net outflow of foreign portfolio investment in 2006.

Critical Issues to Advance Bond Market Development

With a generally benign economic and financial environment and regulatory encouragement, the lack of substantial Arirang issuance must stem from very specific factors. These explicit enabling factors and impediments exist in three areas: supply (issuer), demand (investor), and infrastructure.

Issuer Concerns. Pricing is imperative. The investor market in Korea is one of an increasingly aggressive search for higher yields. Foreign issuers would likely be at the quality end, possibly rated higher than the government's sovereign rating in international markets and close to government securities in domestic markets. Thus pricing remains a difficult issue unless there is local demand for bonds of appropriate risk and maturity. The continuing expansion of the government bond market may also result in considerable crowding-out effects. Pricing at the long end of the yield curve can be problematic, as the pricing benchmark is the three-year treasury bond. Nonetheless, the few issues in 2006 were by quality issuers, which may reflect a broadening in demand by local investors as well as speculators seeking non-U.S. dollar investments.

The Ability to Hedge through Cross-currency Swaps is Imperative. Korea's cross-currency swap market has evolved significantly in the past few years. The interbank trading of foreign currency derivatives alone reached US$3.6 billion daily in 2005, doubling over the previous year and tripling over 2002 (BIS 2005). The Bank of Korea attributes heightened activity—more than 70 percent of it accounted for by swap trades—to greater demand for hedging from trade-oriented local firms. Nonetheless, the derivatives market remains shallow compared with Hong Kong, China, and Singapore, where trading is more than 10 times higher. As a result, pricing is said to be intermittent, and spreads are often wide.
There is a caveat on the importance of derivatives in facilitating bond market development more generally. For example, Singapore has impressive foreign exchange and trading capability, although the Singapore dollar foreign bond market remains small compared with those in the key regional markets of Australia, Hong Kong, and Japan despite recent reforms and initiatives undertaken to develop the local bond market. These include building the government bond market by ensuring a critical mass of investors and issuers, establishing the physical infrastructure, and developing the talent pool (Lian 2002). However, foreign issuers are required to swap or convert proceeds into foreign currency, which in effect prevents unhedged borrowing or the ability of nonresidents to fund Singapore assets.

Clear documentation is important. There is no language barrier for chaebol subsidiaries, but it is considerable for non-Korean firms. Much of the documentation governing the law and legal procedures is largely inaccessible to foreigners. In addition, compliance with issuing procedures has been cited as unduly complex. The requirement to translate foreign documents into Korean is onerous. Also, there have been 15 revisions to the Regulation on Securities Issuance and Disclosure since 2000, and recent revisions are generally not available in English. These are major impediments to nonresident issuance.

The need for domestic ratings is problematic. All foreign issuers must be rated by a local credit rating agency, irrespective of whether they have already been rated by Moody's or Standard & Poor's. This considerably delays issuance. In a recent example, Ford Motor Credit Company was interested in issuing the won equivalent of US$200 million in June 2003. Salomon Smith Barney was selected as the main agency, and Korea Investor Service and National Information & Credit Evaluation were requested to provide ratings. However, documentation and financials were required in Korean, and various participants were unable to act promptly. Meanwhile, the opportunity to achieve sub-LIBOR borrowing rates had changed, so Ford Motor Credit Company did not proceed with the issue (Euroweek, 2003). The issuing procedure can be easier for supranationals, which can classify new issues as government bonds or special bonds. This was critical for the issuing of the Asian Development Bank's first Arirang bond, which served as a benchmark for later issuers. In addition, as Kisselev and Packer (2006) point out, differences of opinion about the creditworthiness of borrowers is especially pronounced among rating agencies in Asia, particularly for local currency ratings. Such discrepancies appear to be more important in other regional markets, where the preference is for quality issuers.

Bond proceeds must not be subject to capital and foreign exchange restrictions. Recent efforts to build financial markets in Korea have been undermined by restrictions on capital movement. The rules are complex, although efforts are being made to
simplify them (IFLR 2005, 2006) as part of a broader government agenda to establish Korea as a regional financial center and, more recently, to internationalize the won. Consider the relative importance of domestic and offshore trading in the Australian Hong Kong, and Singapore dollars and the won. According to BIS (2005; tables E1 and E7), offshore trading accounts for 59 percent of global trading of US$97.0 billion in Australian currency per day; 18 percent of US$33.2 billion in Hong Kong, currency; 19 percent of US$21.2 billion in Singapore currency; and 36 percent of US$17.0 billion in Korean currency. Such evidence suggests the importance of local currency internationalization in developing large foreign bond markets. Singapore is the exception here. The greater internationalization of the Singapore dollar would be expected to be associated with a larger foreign bond market than currently exists. The likely explanation is that the restrictions on bond proceeds have had a negative effect on the size of this market.

**Investor Concerns**

*Market size and liquidity must be enhanced.* The small scope of the Arirang market severely limits its appeal to domestic and foreign investors alike, although these effects may be mitigated through private placement of issued securities. Private placement does not appear to be as important in Korea as in Hong Kong, where almost half the issues are private placements. An interesting feature of the Australian foreign bond market, which is the largest of the markets considered, is that few issues are private placements. Thus, liquidity is assured by the large size of the issue itself (about US$300 million). This suggests that in the absence of a private placement market of sophisticated investors the importance of issue size for inducing higher levels of liquidity cannot be understated. Thus the 55 billion won (US$55 million) average issue size in the Arirang market is likely impeding secondary market trading and liquidity.

*A simple pricing structure may not be the best.* The Arirang market, like the corporate bond market in general, is very homogeneous and does not cater to a broad range of investors. Until recently, most issues have been fixed-rate. This homogeneity is driven in part by the specific nature of local demand. On the other hand, the Arirang bonds of SKC Inc., which include option-like features, demonstrate that there is a market for more complex and tailor-made investment products, most likely as private placements to investors with specific maturity and risk needs, such as insurance companies. What is most likely hindering the development of more varied coupon and pricing features is the absence of developed derivatives markets in floating-rate instruments. The ability of the markets in Australia and Hong Kong, to support issues of floating-rate bonds may be linked to the sophistication of the local two-way swap market (fixed to floating as well as floating to fixed) and forward rate agreements.

Jonathan A. Batten and Peter G. Szilagyi
The won offers limited diversification benefits. Korean bonds offer some scope for diversification since their dollar returns co-move only moderately with U.S. returns (Jiang and McCauley 2004). However, the won itself offers little diversification benefits to most investors, and thus global portfolios stay away from won-denominated debt. Efforts are being made to increase the internationalization of the won, but there is still very little offshore trading in won. Nonetheless, to the extent that the won is perceived to be undervalued against the U.S. dollar, speculators should be encouraged to hold long positions in the currency. However, these positions must also be easy to reverse to reflect changed expectations, and exchange rate regulations tend to dissuade speculative flows. There has been no evidence of speculative foreign investment in the corporate bond market so far, although foreign participation in the three-year government bond futures market is significant.

The withholding tax deters foreign investors. Interest income is subject to a 14 percent withholding tax for resident investors and 25 percent for foreign investors, with reductions or exemptions possible under applicable tax treaties or agreements with the domiciliary country. Recent announcements by the Korean Ministry of Finance and Economy also suggest that withholding taxes will soon be eased to 14 percent (IFLR 2006). Nonetheless, the withholding tax may be a larger barrier than the rates themselves or the bilateral arrangements might suggest. Foreign investors simply do not want to assume the administrative burden of taking advantage of tax treaty rights (Jiang and McCauley 2004). Australia, which offers clear exemptions to withholding taxes, shows that it is possible to encourage international investor involvement in domestic bond markets through appropriate tax policies.

Poor governance and investor protection inhibit foreign investment. In the aftermath of the East Asian crisis, many investors were unable to recover their claims because legal systems inadequately supported investor rights. This problem persists even under normal conditions. The likelihood of bankruptcy filings has been shown to be lower where creditor rights are weaker and the judicial system is less efficient (Claessens, Djankov, and Klapper 2003). Burger and Warnock (2006b) also highlight the importance of investor protection to the development of local bond markets. While creditor protection is reasonably strong in Korea as measured by creditor rights (La Porta and others 1998) and the resolution of claims disputes (Djankov, McLiesh, and Shleifer 2004), these do not compensate for the regulatory and governance inadequacies criticized by the Asian Corporate Governance Association and the World Bank. Investor confidence is also hurt by the aggressive growth strategy of Korean firms such as Hyundai Motor. The effect of these uncertainties is reflected in investor reluctance to hold certain securities and relatively higher credit spreads on corporate bonds.
Infrastructure Issues. Despite improvements, the regulatory and financial infrastructure in Korea is not ready to support a large and sophisticated foreign bond market. Korea ranks only seventh among ten Asian countries—well behind Indonesia and Thailand and below the Philippines as well—on the Hong Kong and Shanghai Banking Corporation's Asian Local Bank Index, which tracks the total return of liquid bonds in local regional economies and assesses the difficulty of accessing bond markets. Korea scores especially low on currency and capital restrictions, the complexity of the withholding tax, and the ease of setting up and operating an investment fund. Korea also receives criticism for infrastructure conditions in pricing, transparency, settlement and custody, derivatives, and hedging.

The technical infrastructure is critical. The Korean government is aware of the need to maintain and lengthen liquidity along the yield curve and to concentrate issuance in specific maturity buckets. New treasury issues are allocated reasonably evenly across three maturities. Three-year bonds are still the most popular, but 5- and 10-year bonds are gaining quickly. While this contributes to better pricing in the corporate bond market, many nonpricing-related issues concern foreign market participants, such as the limited availability of offsetting risk management technologies.

The absence of floating-rate markets is a notable shortcoming. There is also no clear development plan for simple exchange-traded and over-the-counter derivatives. There have been delays in the introduction of interest rate futures and other derivatives linked to benchmark bond curves to facilitate risk management and trading. Several recent initiatives are aimed at enhancing price discovery. One is to increase the use of floating-rate measures, which would eventually lead to the expansion of floating-rate instruments (forward rate agreements and possibly short-dated futures). This process began in 2004 with the introduction of the Korea Interbank Offered Rate (KORIBOR), which should become the benchmark interest rate for short-term financing for banks and may become a reference rate for bond or swap transactions. Hong Kong, Singapore, and Australia have equivalent floating-rate benchmarks. These three markets all have sophisticated forward rate arrangements and futures markets as well.

The Korea Securities Depository has yet to complete full reform of corporate bond trading, settlement, and custody. The foundations of the bond registration system were laid by the Registration Act in 1993. The Korea Securities Depository is finally moving toward the full dematerialization of corporate bonds and is working on introducing electronic trading. It has also allowed listed won-denominated corporate issues to be used in repo transactions, which should increase liquidity in the market.

Capital controls need further attention. Korea still has capital controls (Noland 2005). There are limits on foreign won funding aimed at hedge funds. Foreign sales of
foreign exchange over US$20,000 and purchases over US$10,000 without documentation are subject to notification requirements. The government’s exchange rate policy and currency market interventions go beyond smoothing operations and have been a source of controversy. The effectiveness of capital controls, which led to a buildup of foreign reserves, needs to be reconsidered. Anecdotal evidence suggests that illegal money transfers abroad are both common and considerable, totaling an estimated US$1.2 billion in the first half of 2004.

Reform agenda. The government hopes to secure Seoul’s position as an international financial center by working toward full foreign currency liberalization by 2011. The first phase of the Foreign Exchange Liberalization Plan (2002–2005) increased won funding limits for nonresidents and raised the ceiling on the amount of residents’ foreign borrowings requiring notification. Meanwhile, policy measures are being implemented step by step to deepen and widen the foreign exchange market, such as allowing securities firms into interbank and over-the-counter foreign exchange derivatives trading and forging coalitions between domestic and foreign brokerage firms. By the time the plan is implemented, the government hopes to have internationalized the won and made its foreign exchange market a leading financial hub in the Asia and Pacific region.

The 2005 merger of the Korea Stock Exchange, the Korean Securities Dealers Automated Quotation stock market, and the Korea Futures Exchange into the Korea Exchange is expected to upgrade the competitiveness of the nation’s trading system for a variety of financial products, including stocks, bonds, options, and other derivatives. The Korea Exchange is further easing some restrictions on foreign equity investors, who already account for more than 40 percent of the market. Internationalization of the stock market has done nothing to increase foreign investment in Korean bonds, however. The Korea Exchange seems unclear on how to encourage Arirang issuance other than by easing listing requirements. The cornerstone of its strategy is to scrap the current rule requiring issuers to be listed on either foreign exchanges or have their depository receipts of foreign shares listed on the Korea Exchange as long as the firm is issuing secured bonds with low risk of delinquency.

Lessons and Conclusions

Foreign bond market development is a critical next step for governments in the Asia and Pacific region to achieve the integration of local and global bond markets recommended by McCauley and Park (2006). However, governments with that goal in mind have seldom come up with specific development strategies.
The Korean government, however, considers its plans for reform of the domestic financial system part of an explicit strategy for turning Seoul into an international financial center on a par with Singapore and Hong Kong, by 2010. A critical aspect of this plan is to develop Korea's foreign bond market. The Arirang market has only recently started to attract quality issuers, and its overall size is modest given the scale and scope of the local bond markets.

Funding FDI or local currency portfolios appears to have little to do with why foreign firms enter a domestic bond market. Rather, borrowers seek out minimum cost financing among a spectrum of choices. The appeal of Arirang issuance thus hinges largely on the sophistication of Korea's cross-currency swap market and the ease of transferring these domestic funds to their ultimate destination. The clear lesson from other regional markets, especially Australia, is the importance of swaps and floating-rate derivatives for risk management. As progress is being made on this front, the Korean government must ensure that the other obstacles to easy access to the market are also removed. The tasks include improving pricing conditions, reducing administrative burdens, eliminating capital controls, and encouraging local and international investor demand.

There are numerous lessons for other East Asian governments seeking to develop a foreign bond market. They must create an infrastructure that is enabling for bond issuance and risk management, nurture institutional and retail demand, and deregulate capital markets to facilitate two-way capital flows. There appears to be a natural ordering to the tasks involved: first, establish benchmark bonds and indices; second, develop a diverse derivatives market; third, systematically lengthen the bond market's maturity profile; and fourth, build and develop over-the-counter capability and price structures for derivatives and other complex financial instruments.

Korea's experience also shows that the sophistication of the local bond market does not necessarily make it appealing to foreign borrowers. More is needed.

*Development is Driven by Market Forces and is Difficult to Simply Regulate*

The sophisticated bond markets of Australia, Hong Kong, and Singapore have evolved over many years. An enabling infrastructure is only a first step in market development: the right mix of issuer supply and investor demand is also needed, as Korea shows. Despite its conscious building of infrastructure and the success of its corporate bond market, foreign issuance remains modest. Complex regulatory, issuing, and compliance processes; exchange restrictions; parochial investor attitude; and crowding-out effects may all contribute.
Proper Planning with Industry can Help to Overcome Some Risks

In Australia government withdrawal from the bond markets allowed foreign issuers to substitute at the quality end of the investment spectrum. The reverse appears to be the case in Korea; treasury issuance appears to crowd out foreign borrowers despite low levels of government indebtedness. Filling in infrastructure gaps and maintaining liquidity along the benchmark yield curve in conjunction with industry could overcome some of these problems.

An Order of Issuance Helps Build Markets, but Follow-through is Needed

In Korea corporate borrowers were not granted access to the Arirang market until four years after the first supranational issue. The mandate of supranational institutions is to “provide cost-effective funding on a sustainable basis for financing economic development.”¹¹ They are therefore driven to minimize funding costs and lack the ability to cross-subsidize market development through regular issuance. While they may lead the way with a foreign bond issue, as advocated by Hoschka (2005), there should be plans for follow-up issues by other issuers to maximize liquidity and maintain interest in the market. Such a strategy requires ongoing consultation between industry and government.

Foreign Issuers are Driven by Cost and are not Necessarily Interested in using Local Issuance to Fund FDI or Portfolio Investment

Multinational firms pick financing arrangements very carefully. Chaebol subsidiaries issue Arirang bonds to minimize the funding costs of foreign investment. Chaebols obviously have a comparative advantage at home, which they are able to translate into lower costs in the host country. It is unsurprising that chaebol subsidiaries in Indonesia are at the forefront of the Arirang market, as they would probably have difficulty accessing local debt and banking markets.

Foreign bond markets emerge when local market conditions are highly evolved. It is thus striking that the sophisticated Korean bond market developed with minimal foreign involvement. The potential benefits of a well-developed Arirang market are vast. It should help foreign borrowers to currency-match foreign assets or simply provide an alternative source of funds that can be swapped into the currency of choice as windows of opportunity appear. Foreign bonds could ultimately be an important vehicle to promote cross-border investment, while also making the local Korean bond market larger, more liquid, and more resilient to boom and bust cycles. While there is no systematic or institutional discrimination against foreign bond issues, current conditions discourage cross-border activity, amplified by cumbersome access, governance concerns, and weaknesses in
institutional arrangements. There appears to be no basis to concerns that expansion of the Arirang market will crowd out local offerings or have destabilizing exchange rate effects.

Korea's experience can provide insights for the ongoing reform of developing economies in Eastern Europe and Asia. In developed countries, as banks have become increasingly cautious about extending credit, a gradual process of disintermediation has been occurring in historically bank-oriented financial regimes, fed by considerable regulatory efforts directed at market liberalization. In emerging market economies bank-intermediated finance is the single most important source of funds to the corporate sector. As these economies experience the same trend toward disintermediation they urgently need to open new financial channels. Korea's example, with all its imperfections, can provide a helpful guide.

### Appendix

**Table A1. Arirang Bond Issues, June 1995–December 2006**

<table>
<thead>
<tr>
<th>Issuer name</th>
<th>Domicile of issuer</th>
<th>Local credit rating</th>
<th>Issue date (DD/MM/YYYY)</th>
<th>Description (%)</th>
<th>Amount (million) Won</th>
</tr>
</thead>
</table>
| Asian Development Banka Supranational | AAA | 01/09/1995 | 7-year fixed 12.15 | 80,000
| European Bank for Reconstruction and Development (EBRD)a | Supranational | AAA | 23/04/1997 | 5-year fixed 9.8 | 71,500
| EBRDa | Supranational | AAA | 02/05/1997 | 5-year fixed 10 | 71,100
| PT CS Indonesiaa | Indonesia | AAA | 11/06/1999 | 3-year fixed 7 | 45,000
| PT Hanil Jayaa | Indonesia | A | 16/07/1999 | 3-year fixed 7 | 14,000
| PT SK Keraa | Indonesia | A | 29/11/1999 | 1-year fixed 8 | 45,000
| PT Indomiwon Citraa | Indonesia | A | 20/12/1999 | 1-year fixed 10.8 | 20,000
| PT Miwon Indonesia TBKa | Indonesia | A | 16/03/2000 | 1-year fixed 9 | 20,000
| PT Indomiwon Citraa | Indonesia | BBB- | 20/12/2000 | 2-year fixed 10 | 10,000
| PT Indomiwon Citraa | Indonesia | BBB- | 28/02/2001 | 1-year fixed 11.13 | 15,000
| PT Miwon Indonesia TBKa | Indonesia | BBB- | 16/03/2001 | 1-year fixed 10.99 | 20,000
| SKC Inc.a United States | BBB | 03/05/2001 | 2-year fixed 10.2 | 50,000
| PT SK Keraa | Indonesia | BBB | 28/05/2001 | 3-year fixed 9.875 | 30,000
| Hyundai Assan OSVTa | Turkey | A- | 04/06/2001 | 3-year fixed 7 | 29,000
| SKC Inc.a United States | BBB | 19/07/2001 | 2-year fixed 8.665 | 70,000
| PT SK Keraa | Indonesia | BBB | 23/07/2001 | 3-year fixed 8.85 | 50,000
| SKC Inc.a United States | BBB | 06/11/2001 | 2-year fixed 8.27 | 50,000
| SPI (Seosan) Cogena Singapore | A | 20/12/2001 | 3-year fixed 7.68 | 140,000
| SPI (Seosan) Watera | Singapore | A | 20/12/2001 | 5-year fixed 7.68 | 70,000
| PT Indomiwon Citraa | Indonesia | BBB | 27/02/2002 | 1-year fixed 7.26 | 16,000

*Continued*
Table A1. Continued

<table>
<thead>
<tr>
<th>Issuer name</th>
<th>Domicile of issuer</th>
<th>Local credit rating</th>
<th>Issue date (DD/MM/YYYY)</th>
<th>Description (%)</th>
<th>Amount (million) Won</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT CS Indonesia(^b)</td>
<td>Indonesia</td>
<td>A +</td>
<td>27/04/2002</td>
<td>1-year fixed 7.32</td>
<td>100,000</td>
</tr>
<tr>
<td>PT Indomiyon Citra</td>
<td>Indonesia</td>
<td>BBB</td>
<td>17/10/2002</td>
<td>3-year fixed 6.43</td>
<td>30,000</td>
</tr>
<tr>
<td>SKC Inc.(^a)</td>
<td>United States</td>
<td>BBB +</td>
<td>02/05/2003</td>
<td>1-year fixed 7.75</td>
<td>50,000</td>
</tr>
<tr>
<td>SKC Inc.(^a)</td>
<td>United States</td>
<td>BBB +</td>
<td>21/07/2003</td>
<td>1-year fixed 10.02</td>
<td>70,000</td>
</tr>
<tr>
<td>PT SK Keris(^a)</td>
<td>Indonesia</td>
<td>BBB +</td>
<td>22/07/2003</td>
<td>2-year fixed 9.2</td>
<td>50,000</td>
</tr>
<tr>
<td>SKC Inc. (^ab)</td>
<td>United States</td>
<td>BBB +</td>
<td>06/10/2003</td>
<td>3-year fixed 9.5</td>
<td>5,000</td>
</tr>
<tr>
<td>SKC Inc. (^ab)</td>
<td>United States</td>
<td>BBB +</td>
<td>20/10/2003</td>
<td>3-year fixed 9.5</td>
<td>10,000</td>
</tr>
<tr>
<td>SKC Inc.</td>
<td>United States</td>
<td>BBB +</td>
<td>06/11/2003</td>
<td>1-year fixed 7.24</td>
<td>35,000</td>
</tr>
<tr>
<td>SKC Inc.</td>
<td>United States</td>
<td>BBB +</td>
<td>03/05/2004</td>
<td>1-year fixed 6.1</td>
<td>50,000</td>
</tr>
<tr>
<td>Standard Chartered PLC</td>
<td>Great Britain</td>
<td>AAA</td>
<td>27/04/2004</td>
<td>10-year floating</td>
<td>50,000</td>
</tr>
<tr>
<td>SKC Inc.</td>
<td>United States</td>
<td>BBB +</td>
<td>21/07/2004</td>
<td>1-year fixed 6</td>
<td>45,000</td>
</tr>
<tr>
<td>SPI (Seosan) Water</td>
<td>Singapore</td>
<td>A +</td>
<td>01/12/2004</td>
<td>3-year fixed</td>
<td>50,000</td>
</tr>
<tr>
<td>SKC Inc.</td>
<td>United States</td>
<td>BBB +</td>
<td>08/11/2004</td>
<td>2-year fixed 5.1</td>
<td>30,000</td>
</tr>
<tr>
<td>SPI (Seosan) Cogen</td>
<td>Singapore</td>
<td>A +</td>
<td>01/12/2004</td>
<td>3-year fixed</td>
<td>100,000</td>
</tr>
<tr>
<td>PT SK Keris</td>
<td>Indonesia</td>
<td>AAA</td>
<td>06/12/2004</td>
<td>3-year fixed 5.0</td>
<td>40,000</td>
</tr>
<tr>
<td>PT Chiel Jedang Indonesia</td>
<td>Indonesia</td>
<td>AA-</td>
<td>30/03/2005</td>
<td>3-year fixed 4.72</td>
<td>100,000</td>
</tr>
<tr>
<td>PT Chiel Jedang Indonesia</td>
<td>Indonesia</td>
<td>AA-</td>
<td>30/03/2005</td>
<td>3-year fixed 4.72</td>
<td>100,000</td>
</tr>
<tr>
<td>SKC Inc.</td>
<td>United States</td>
<td>BBB +</td>
<td>03/05/2005</td>
<td>2-year fixed 4.77</td>
<td>30,000</td>
</tr>
<tr>
<td>PT SK Keris</td>
<td>Indonesia</td>
<td>BBB +</td>
<td>21/07/2005</td>
<td>3-year fixed 5.5%</td>
<td>50,000</td>
</tr>
<tr>
<td>SKC Inc.</td>
<td>United States</td>
<td>BBB +</td>
<td>21/07/2005</td>
<td>2-year fixed 4.91</td>
<td>45,000</td>
</tr>
<tr>
<td>NAC Funding Ltd</td>
<td>Ill</td>
<td></td>
<td>26/07/2005</td>
<td>5-year fixed 4.92</td>
<td>120,000</td>
</tr>
<tr>
<td>PT SK Keris</td>
<td>Indonesia</td>
<td>BBB +</td>
<td>30/11/2005</td>
<td>2-year fixed 6.13</td>
<td>20,000</td>
</tr>
<tr>
<td>SKC Inc.</td>
<td>United States</td>
<td>BBB +</td>
<td>26/12/2005</td>
<td>2-year fixed 6.3</td>
<td>30,000</td>
</tr>
<tr>
<td>SBC Corp BHD</td>
<td>Malaysia</td>
<td>AAA</td>
<td>28/6/2006</td>
<td>8-year fixed 5.2</td>
<td>100,000</td>
</tr>
<tr>
<td>SBC Corp BHD</td>
<td>Malaysia</td>
<td>AAA</td>
<td>28/6/2006</td>
<td>8-year fixed 5.2</td>
<td>100,000</td>
</tr>
<tr>
<td>Citicorp Capital Markets Australia LTD</td>
<td>Australia</td>
<td>AA</td>
<td>28/6/2006</td>
<td>3-year fixed 5.07</td>
<td>20,000</td>
</tr>
<tr>
<td>Citicorp Capital Markets Australia LTD</td>
<td>Australia</td>
<td>AA</td>
<td>28/7/2006</td>
<td>3-year fixed 5.07</td>
<td>20,000</td>
</tr>
<tr>
<td>Citicorp Capital Markets Australia LTD</td>
<td>Australia</td>
<td>AA</td>
<td>25/9/2006</td>
<td>2-year fixed 4.96</td>
<td>10,000</td>
</tr>
<tr>
<td>Citicorp Capital Markets Australia LTD</td>
<td>Australia</td>
<td>AA</td>
<td>25/9/2006</td>
<td>2-year fixed 4.96</td>
<td>10,000</td>
</tr>
</tbody>
</table>

| Total issues | 2,486.600 |
| Average (49) issue size | 50,746 |
| Average maturity | 3-years |

| Bear Stearns Co Inc. | United States | A1 | 26/6/2006 | 7-year floating | 200 |
| Bear Stearns Co Inc. | United States | A1 | 26/6/2006 | 7-year floating | 200 |
| PT Chiel Jedang Indonesia | Indonesia | 9/10/2006 | 3-year floating | 30 |

Continued
Table A1. Continued

<table>
<thead>
<tr>
<th>Issuer name</th>
<th>Domicile of issuer</th>
<th>Local credit rating</th>
<th>Issue date (DD/MM/YYYY)</th>
<th>Description (%)</th>
<th>Amount (million)</th>
<th>Won</th>
</tr>
</thead>
<tbody>
<tr>
<td>PT Cheil Jedang Indonesia</td>
<td>Indonesia</td>
<td></td>
<td>9/10/2006</td>
<td>3-year floating</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>NAC Funding PLC</td>
<td>Ireland</td>
<td></td>
<td>6/1/2006</td>
<td>5-year fixed</td>
<td>5.17</td>
<td>40</td>
</tr>
<tr>
<td>NAC Funding PLC</td>
<td>Ireland</td>
<td></td>
<td>6/1/2006</td>
<td>5-year fixed</td>
<td>5.17</td>
<td>40</td>
</tr>
</tbody>
</table>

*Matured.
*Callable.
*Private placement.
Source: Reuters Fixed Income Database.

Notes

Jonathan A. Batten (corresponding author) is professor of finance in the Graduate School of Management at Macquarie University, Sydney, and adjunct professor at Hong Kong University of Science and Technology; his email address is jabatten@gmail.com. Peter G. Szilagyi is a lecturer in finance at the Judge Business School, University of Cambridge; his email address is p.szilagyi@jbs.cam.ac.uk. The authors are grateful to Barry Eichengreen, Tobias Hoschika, Patricia McKean, Jim Turnbull, Richard Werner, and three anonymous referees for helpful comments; Reuters Asia Pte Ltd for bond data; and the Asian Development Bank for providing funding to support this research. Preliminary findings of this research were presented at the ASEAN + 3 Deputies Meeting in November 2004.

1. *Asiamoney* (2006) notes that despite the recent reform efforts of East Asian countries, 70 percent of cross-border portfolio investment originating in the region remains fixed rate and 80 percent is directed to the United States and Europe. Strikingly, only 5 percent is directed to other emerging markets in Asia, whereas 63 percent of European cross-border investment stays in Europe.

2. See IMF (2005) for further discussion.

3. Note the JP Morgan Emerging Markets Bond Index dropped from 800 basis points above U.S. Treasuries in 2001 to 200 basis points at the end of 2006.

4. As Eichengreen and Luengnaruemitchai (2004) note, outstanding debt securities in Asia are on average 9.27 percent of GDP for corporate bonds, 23.52 percent of GDP for public sector bonds, and 12.0 percent of GDP for financial institutions. This compares with 20.55 percent, 85.15 percent, and 33.64 percent for developed countries.


6. On the May 19, 2006, the Korean Ministry of Finance and Economy accelerated its schedule for liberalization of the won and capital flows (McCauley 2006).

7. These themes are further developed by Rhee (2004), who discusses how a regional bond market facilitates credit enhancement and corporate risk management.

8. See Woo (2002) for a more detailed discussion of the historical development of the Korean bond market.

9. *Chaebol* is the Korean term for conglomerate, such as the Samsung Corporation and its many affiliates. A key feature of these corporations is their complicated ownership structures, through cross-ownerships of shares with different types of voting rights that frequently allow control to remain with the founding family.

Jonathan A. Batten and Peter G. Szilagyi
10. The authors thank an anonymous referee for bringing this point to our attention.


12. Consider 2006 issues in the Malaysian ringgit (M$) market: KIW priced its M$500 million seven-year issue with a coupon of 4.6 percent, which is 46 basis points over the benchmark five-year Malaysian government security, while the Asian Development Bank priced its five-year issue at 15 basis points over Malaysian government securities. Both were swapped into US dollars to achieve sub-LIBOR financing.

13. For example, the Korean Financial Supervisory Service (FSS) notes the following disclaimer: “The English translation of the financial supervisory regulations is not official and is intended for reference only. Neither the FCC [Financial Supervisory Commission] nor the FSS is responsible for the correctness of the English translation, and the reader is advised to refer to the most up-to-date regulations in Korean. The English translation is current as of August 1, 2002.” See http://english.fss.or.kr/en/laws/sec/lawstock.jsp


15. This is a quotation from the World Bank website on their funding objectives, although it is consistent with others. See http://treasury.worldbank.org/Services/Capital%2bMarkets/Annual+Issuance/Funding+Strategy+and+Objectives.html

References


Inside Decentralization: How Three Central American School-based Management Reforms Affect Student Learning Through Teacher Incentives

Ilana Umansky • Emiliana Vegas

Despite decentralization reforms of education systems worldwide, there is little empirical evidence about the processes through which decentralization can improve student learning. Proponents theorize that devolving decisionmaking authority to the local level can improve communication, transparency, and accountability, making teachers and school principals more responsible for better performance and more capable of bringing it about. Yet some research has shown that decentralization can increase inequality and reduce learning for disadvantaged students. This article reports on retrospective evaluations of three Central American school-based management reforms. Using matching techniques, these evaluations investigate whether the reforms enhanced student learning and how they affected management processes and teacher characteristics and behaviors. The evidence indicates that all three reforms resulted in substantive changes in management and teacher characteristics and behavior and that these changes explain significant portions of resultant changes in student learning. This article contributes to the understanding of how decentralization reforms can improve learning and shows how education reforms, even when not conceptualized as affecting teacher incentives, can generate important changes for teachers that, in turn, affect student learning. JEL codes: I21, 128, H52, H75.

In countries around the world, governments are decentralizing the management of public services. Educational decentralization, often in the form of school-based management, is widespread in Africa, Asia, and Latin America. Many of these reforms sprang from the notion that schools can better meet the needs of students
and communities by responding and being directly accountable to local stakeholders. The goals of educational decentralization frequently include improved education quality, greater relevance, expanded access, and increased efficiency. But exactly how educational decentralization will achieve these wide-ranging improvements has been left largely at the level of broad theoretical assumptions.

This article examines whether school-based management improves student learning. It investigates how school-based management reforms influence teachers and how that influence affects student outcomes. The goal is to broaden the conception of how education reforms affect teachers by influencing teacher incentives and to explore how to design and implement these reforms to maximize their beneficial effects on teaching and learning.

Three school-based management reforms in Central America are analyzed: Education with Community Participation (EDUCO) in El Salvador, the Honduran Community Education Project (PROHECO) and School Autonomy in Nicaragua. These three reforms have altered the incentives teachers face, changing teacher composition and behavior, with direct implications for student learning. This article shows that education reform design should consider the potential impact on teaching quality, even when reforms are not specifically intended to alter the incentives that teachers face.

This article draws on three background papers (Sawada and Ragatz 2005; di Gropello and Marshall 2005; Parker 2005) prepared for a regional study of teacher quality and incentives in Latin America published by the World Bank (Vegas 2005).

The article first looks at the literature on how school-based management reforms affect teachers and then proposes a theoretical framework for understanding these reforms as teacher incentive reforms. Next, it provides background information on the three Central American school-based management reforms and reviews the data and methodology in the analyses of these reforms. The following sections explore how the reforms affected teachers and students. The final section discusses policy implications and conclusions for designing and implementing reforms that can benefit teaching and learning.

The Literature on School-based Management Reforms and Teachers

Educational decentralization reforms have varied widely in their content, goals, areas of decisionmaking, and levels of educational responsibility (McGinn and Welsh 1999). Empirical analyses have also shown a wide range of impacts on education quality. Some studies have established a clear relationship between
educational decentralization and improved indicators of student learning, such as test scores, repetition rates, and dropout rates (Stoddard and Kuhn 2006; Lavy 2004; Galiani and Schargrodsky 2001; Eskeland and Filmer 2002; Jimenez and Sawada 2003). Other evidence, however, points to the importance of contextual issues for the effectiveness of educational decentralization reforms in improving learning outcomes. In the wrong context, educational decentralization can even diminish learning and exacerbate inequalities in learning outcomes (Gunnarson and others 2004; Galiani, Gertler, and Schargrodsky 2005; Hsieh and Urquiola 2003; Fuller and Rivarola 1998; McGinn and Welsh 1999).

Few studies have investigated the impact of decentralization on teachers and their work. Winkler and Gershberg (2000) propose a theoretical model for the ways decentralization could stimulate skilled and committed teachers. With localized resources, schools or local governments can make more appropriate decisions about professional development and training for teachers. With localized evaluation, school directors can provide feedback on teachers' work and how to improve it. Greater teacher autonomy in terms of curricula, pedagogy, or school improvement plans can motivate and enable teachers to improve learning in their schools.

Empirical research confirms some of these possible relationships. In a study of the Chilean voucher system, Vegas (2002) finds evidence that certain management strategies in schools are associated with higher student performance when controlling for student socioeconomic background. She finds that greater teacher autonomy in implementing projects and designing teaching plans is associated with better student outcomes when school decisionmaking power is close to the level of the teacher. In parallel with the work of Heneman (1997), this finding likely indicates that teachers are better able to translate greater autonomy into improved teaching if they have support from their supervisors.

Jimenez and Sawada (2003) demonstrate that direct community management of schools can improve teacher effort. They find that teachers in El Salvador's community-managed EDU CO schools exert significantly greater effort than teachers in centrally managed schools. The authors hypothesize that community control over the school budget allows community members to exert meaningful pressure on teachers and to design compensation systems that motivate greater teacher effort. This local accountability, they argue, is in turn the basis for higher student academic outcomes in EDU CO schools than in traditional schools.

Evidence from the United States, however, shows that accountability and school choice reforms have not resulted in greater teacher effort, as measured by weekly hours worked. For example, Stoddard and Kuhn (2006) find that average weekly work hours of U.S. teachers has increased steadily in both states with and without accountability and school choice reforms. They hypothesize that the
weak link between teacher effort and compensation may explain the lack of greater teacher effort as a result of the reforms.

In contrast to Winkler and Gershberg’s (2000) theory of greater teacher autonomy under decentralization, evidence from the Nicaraguan School Autonomy reform suggests diminished autonomy for teachers participating in the reform (King and others 1996). Teacher survey data show that teachers in autonomous schools report less influence over school decisions than they had in traditional public schools. However, teachers also report that they are more punctual and less frequently absent in autonomous schools than they were in traditional schools.

In a qualitative study of 12 autonomous schools in Nicaragua, Fuller and Rivarola (1998) indicate that teachers’ financial difficulties associated with the reform may have negative ramifications for teaching and learning. They find that teachers in autonomous schools spend considerable instructional time on collecting fees from students and have begun to offer more out-of-class tutorials for which children must pay.

The studies of decentralized and community-managed schools demonstrate the diversity in decentralization reforms, as well as the powerful implications of seemingly small details of planning and implementation. Whereas in El Salvador local accountability to parents seems to be improving teacher effort, in Nicaragua evidence suggests that user fees are draining classroom instructional time and teacher morale.

How Reforms Can Affect Teacher Characteristics and Behavior

This section develops a theoretical framework for understanding how school-based management reforms are likely to generate teacher incentives and what kinds of impact these incentives may have on teacher characteristics and behavior. It considers both how reforms might result in more skilled teachers and teaching as well as ways in which the reforms might result in lower quality teachers and teaching.

A useful point of departure for understanding the impact on teachers is principal-agent theory. Principals, such as employers, design compensation structures to get agents, such as employees, to work in the principals’ interest (Ross 1973). In education, the principal–agent relationship can take multiple forms. Teachers, as agents, can be considered to be working on behalf of multiple principals, including parents, school principals, and education officials. In the case of school-based management reforms, the principal is generally the school council. Principal–agent theory assumes that the interests of principals and agents are frequently not aligned. Principals want high employee productivity and efficiency,
while agents want high compensation for little effort. Principals seek means of motivating agents to behave in ways that principals believe will result in high productivity and efficiency.

To what extent agents alter their behavior depends partly on their degree of risk aversion, their assessment of the risk involved in the behavior, and the desirability of the reward or aversion to the sanction (Baker 2002; Prendergast 1999). A potentially large reward or large sanction is more likely than a small reward or sanction to elicit substantial effort by a teacher to change behavior. Likewise, a reward or sanction for a change in behavior that involves little effort or risk is more likely to elicit the desired change than a reward for behavior that requires substantial effort or involves significant risk.

School-based management has the potential to generate several incentives and conditions that can affect teacher quality and teaching. These include greater accountability to local stakeholders, direct communication between communities and schools concerning their needs and interests, more flexible and meritocratic pay and employment, advancement structures associated with closer-to-the-source evaluation, school-based employment decisions, and weaker teachers unions. School-based management can make teachers more accountable to the needs and interests of the families of their students, particularly when local school boards or parent associations have the authority to hire and fire teachers or set teachers' wages. In theory, this very direct control over teachers' work generates strong incentives for teachers to satisfy community interests and could improve teachers' work and expand student learning.

In addition, having local school boards consisting of parents, community members, and students can improve the lines of communication between teachers, school managers, and administrators and the communities they serve. In centralized education systems, communication is indirect, often filtering imperfectly from communities to education ministries and then back to principals and teachers. The direct and localized communication in locally managed schools could also result in improved teaching and learning, particularly if the centralized system is inefficient and ineffective.

School-based management is also generally associated with weaker teachers unions, as teachers in these schools are often subject to local decisions rather than national guidelines established through collective bargaining. Previous research has indicated that weaker teachers unions are sometimes associated with improved teacher quality and teaching because in countries with strong teachers unions teachers' positions are typically based on rigid pay and advancement structures, such as seniority and education level, rather than on quality of work.

Incentives that could result in better teacher quality and teaching include both "stick" incentives and "carrot" incentives. If teachers perceive local management as disempowering and increasing their vulnerability to local stakeholders, then
school-based management could act as a stick incentive, with teachers working harder for fear of being fired. If, however, decentralization of authority gives teachers more autonomy and ownership in their schools, school-based management could create carrot incentives. Such incentives could also affect teachers in multiple ways, for example, by increasing their enjoyment of their work and their sense of making a difference in children’s lives.

An alternate hypothesis is that school-based management does not, in most instances, improve teachers and their work. Weakened teachers unions and centralized guidelines may destabilize teachers’ work, increasingly disempowering teachers and diminishing the quality of their teaching. Weaker teachers unions could also result in lower salaries and nonmonetary compensation, such as pensions and other benefits. The upshot is that teachers are more likely to abandon their jobs, and highly skilled individuals are less likely to enter teaching.

Locally managed schools in communities that lack the ability to manage their schools are susceptible to increased mismanagement and corruption. These schools may not receive the support that they need from centralized authorities, and they could end up with insufficient resources, training, and general support for teachers.

Central America’s School-based Management Reforms and Their Evaluations

El Salvador was the first Central American country to experiment with school-based management reform. In the 1980s, many poor rural communities were cut off from central services because of the civil war. Lacking functioning schools, communities decided to create and run their own local schools. After the war ended, the central government acknowledged the success of these schools in providing education cost-effectively in remote areas and in 1991 decided to expand the program. Schools in the EDUCO program are managed by community associations with block grants from the central government.

In 1999, Honduras launched a similar reform, PROHECO, aimed at expanding and improving cost-effective, community-run primary and preprimary schools in isolated rural areas of the country. In both El Salvador and Honduras most communities without schools and with a minimum number of school-age children can receive government funds to start a locally managed school.

The reform in Nicaragua was different. Piloted in 1993, Nicaragua’s School Autonomy program was aimed initially at urban secondary schools, in particular at existing schools with higher than average resources. As it has grown, the School Autonomy reform has incorporated many primary schools and rural schools.
By 2002, 63 percent of Nicaraguan students were in autonomous schools, and 37 percent of primary and secondary schools were autonomous schools.

All three reforms have in common the establishment of school councils, made up of community and school representatives, as legal entities with decisionmaking authority. The councils have the power to hire teachers and maintain school buildings. EDUCA also gave the council the authority to evaluate teachers and to build the local school. PROHECO decentralized control over selecting teachers, monitoring student and teacher performance, managing funds for school resources and improvement activities, and building the local school. The Nicaraguan reform gave councils the authority to hire administrative staff, manage the school budget, raise revenues, evaluate teacher performance, and make some pedagogical choices (di Gropello 2006).

The evaluations of the impact of these reforms on teaching quality and student outcomes used quasi-experimental techniques. In all countries, only a portion of schools participated in the reform. Because the reforms were not implemented under experimental conditions, with preselected treatment and control schools, the authors of the studies used matching techniques to establish control groups. Since participating schools may differ in systematic ways from other schools, the authors used econometric techniques to address problems of selection bias in constructing comparison groups.

The objective of these evaluations was to estimate the impact of the school-based management intervention, $I$, on results of interest, $Y$, such as characteristics of the teachers, specific teacher behaviors, and student learning outcomes. In general, $Y$ is determined not only by the program (if at all), but also by observed characteristics of the teacher, the school, and the community, $X$; by a host of factors also pertaining to the teacher, the school, and the community that are unobserved, $Z$; and by an error term, $e$, that captures any factors not included in $I$, $X$, or $Z$. This relationship can be written as a linear function.

\[
Y = aI + bX + cZ + e
\]  

where each variable and its coefficient estimate are vectors. Each variable is indexed by its level of aggregation (such as the teacher, school, or community) and time period, but for simplicity of exposition the subscripts have been omitted. The error or disturbance term represents stochastic or chance events that are independent of the right-side variables ($I$, $X$, $Z$) in relation (1).

The main coefficient of interest is $a$. If the reform being evaluated were randomly assigned or placed, it would be possible to make a simple comparison of outcomes $Y$ with and without the reform to get this estimate. Generally, outside of an experiment, a good estimate of $a$ cannot be obtained simply by comparing the mean values of $Y$ with and without the reform or by using simple standard estimation methods such as ordinary least squares. Several estimation problems
arise. One is that the reform being evaluated may itself reflect choices such that the without-the-reform state (or control or comparison group) is not the appropriate counterfactual.

If a school's choice or assignment into the program is based on expected benefits from participation, such as a school's need or likelihood of success, as is almost certainly the case in the Central American reforms, then the characteristics of those who participate will not be comparable to the characteristics of those who do not. When an experimental method of evaluation is not possible, several quasi-experimental evaluation methods can be used (Grossman 1994; Heckman and Smith 1995; Ravallion 2001). These include evaluation of participating schools before and after the reform; matched comparison, in which a comparison group is chosen to match the observed characteristics of the treatment group; propensity score matching, which involves constructing a comparison group based on its conditional probability of receiving treatment given a set of observable characteristics; and natural experiments, which are naturally occurring experimental conditions due to quirks, isolated changes, or idiosyncrasies in the identification of the treatment group.

For the three Central American reforms, the authors constructed ex post comparison or control groups. The authors then estimated the impact of school-based management on teaching and learning using a two-step process. First, they analyzed how characteristics of teachers and of the teaching and learning environment differ between the treatment and control groups. These are considered the first-order results. Because of data limitations, the three studies were not able to investigate these first-order effects to the same extent. Because the groups were selected to be comparable, however, the differences between the groups can be assumed to be the result of the school-based management reform. In a second step, the studies then looked at how these differences in teacher characteristics and teaching and learning environment are associated with higher or lower student performance. The analyses also looked at how simply being in the treatment group is related to student performance. These are considered the second-order results.

This two-step process enables determining whether the school-based management reforms are or are not associated with improved or diminished student learning and whether any differences in student learning are likely to be the result of specific changes in teachers and the teaching environment that resulted from the reforms. The causal nature of the relationships reported here are predicated on accurate data, appropriate methodologies, and the absence of sample-selection bias. The authors of all three background papers addressed these issues, in particular employing techniques to eliminate or test for the presence of sample-selection bias. Thus, the case for causality appears to be reasonably good.
The El Salvador background study uses both matched comparison and propensity score matching techniques to establish control groups (Sawada and Ragatz 2005). The condition of the EDUCO schools tends to be worse—in terms of basic utilities and resources—than that of the control group of centrally managed schools, suggesting that there may be a downward bias, if any, in the treatment effect estimates. Under such conditions Type I errors (false positives) are unlikely. The EDUCO study relies on school survey data from 1996 (more recent survey data were unavailable), national exam results, and municipality-level socio-economic indicators. First, the authors examine how EDUCO schools differ from traditional rural schools in teacher characteristics and behaviors as well as teacher management and school administration. Second, they look at how these factors are associated with student test scores in both the EDUCO and the traditional rural schools.

The Honduras study uses data from 2002 and 2003 school surveys and 2003 student test results (di Gropello and Marshall 2005). The treatment group data comes from the 2003 survey, while two different control groups are used, one drawn from data from the same 2003 survey and one from a 2002 survey. Overall, the authors find consistency in the results of the multivariate analyses between the two control groups. The authors use an instrumental variable analysis—which substitutes a modeled likelihood of being a PROHECO school for whether the school was actually a PROHECO school—to control for sample selection bias.

The study of School Autonomy in Nicaragua uses 2002 survey data and Spanish and math test data for third and sixth graders for both the treatment group and the control group (Parker 2005). Two propensity score matching techniques, nearest neighbor matching, and stratification matching are used to control for selection bias. The reform originally targeted schools with higher school and student resource levels. Because there is generally a direct relation between level of school and student background resources and student performance, this case more than the other two risks an upward bias in estimated impact coefficients. Notably, however, this also means that Type II errors (false negatives) are unlikely.

First-order Effects: Implications of School-based Management Reforms for Teachers

Analysis of first-order results reveals that the school-based management reforms are associated with several important changes in school management and teacher characteristics and behaviors but that these changes are inconsistent across countries. The reforms affect both who works as a teacher and the choices
and behavior of teachers as they teach. Although different in each country, school-based management appears to be causally linked to differences in teachers' education level and years of experience, number of days teachers are absent, number of hours a week they work, amount of homework they assign, teaching methods they use, and general attitude toward teaching.

Although seemingly counterintuitive, the three Central American cases suggest that schools that participate in the reforms do not always have more autonomy than traditional schools. Once other factors are controlled for, in only 3 or 4 of 10 areas do local stakeholders report significantly different degrees of school-level control over administrative processes in El Salvador's EDUCO schools than in traditional schools (table 1).

The areas in which local stakeholders report significantly greater influence in EDUCO schools compared with traditional schools are: hiring and firing administrators, hiring and firing school directors, and hiring and firing teachers. Although the limited areas of greater local management indicate a muted potential impact on teachers, the authority to hire and fire, and to hire and fire teachers, in particular, could potentially have a strong impact on the quality of teaching. Local hiring and firing could hold teachers more accountable to local stakeholders, such as the school principal, school council members, and parents, and it could motivate teachers to work harder or in different ways because of weaker job security or new or different criteria for evaluation.

Decentralization reforms also appear to alter the distribution of authority across local stakeholders. In all three cases, school-based management is designed to increase the involvement and decisionmaking power of parents. In El Salvador,
EDUCO parent associations report high levels of influence relative to traditionally managed schools. Of 29 administrative processes EDUCO parent association members were more likely to report greater control than their non-EDUCO counterparts in all areas, and these results were statistically significant in 26 of the 29 processes. In contrast, school directors in EDUCO schools report less influence than do directors in traditional schools, although only one of the 29 processes was statistically significant (hiring and firing of director). The results for teachers were not clear, as they were statistically significant in the ordinary least squares analysis but insignificant in the propensity score analysis.

In Honduras, the results are similar. PROHECO school principals report significantly less autonomy in their work than principals in traditional schools, and PROHECO teachers are significantly less likely to report a sense of control over school priorities than are teachers in traditional schools. As in the EDUCO reform, local authority appears to have shifted from school directors toward parents, as intended.

In Nicaragua, previous research has concluded that as a result of the reform much of the localized power has gone directly to the school principal rather than to school boards. Parent associations and teachers report little decisionmaking power in autonomous schools. Previous research also established that much of the decentralized authority was administrative or financial rather than pedagogical (Castillo 1998; Fuller and Rivarola 1998; King and others 1996).

Alongside these changes in school management, the reforms have resulted in changes in teacher characteristics and behavior. Table 2 summarizes some of the statistically significant findings regarding these changes. In both El Salvador and Honduras, decentralized schools are closed less frequently than are traditional schools. Evidence from El Salvador suggests that this is due largely to fewer organized labor stoppages. In El Salvador, teachers in decentralized schools are also less frequently absent. There is some evidence in Honduras that teachers in decentralized schools may also be absent less frequently because of less union participation, although they are more likely to be absent because of teacher professional development activity. Because more instructional time is consistently found to improve student learning, it is probable that these characteristics are an improvement over traditional schools. Although the relationship was not investigated econometrically, these changes in teacher behavior are likely in part to be the result of more localized management.

At the same time union-instigated stoppages are usually due to important concerns among teachers regarding pay and working conditions. Indeed, PROHECO teachers in Honduras are more likely to report being unsatisfied with their salary and are, on average, paid less than teachers at traditional schools. These differences may have a negative impact on willingness to remain in teaching, teacher motivation, or other important qualities of effective teaching.
Table 2. Statistically Significant Findings on Teacher Characteristics and Behavior in School-based Management Reform Schools Compared with Traditional Schools

<table>
<thead>
<tr>
<th>Teacher characteristics and behavior</th>
<th>El Salvador EDUCO</th>
<th>Honduras PROHECO</th>
<th>Nicaragua School Autonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational attainment</td>
<td>More</td>
<td>Less</td>
<td>No difference</td>
</tr>
<tr>
<td>Years of experience</td>
<td>Less</td>
<td>Less</td>
<td>No difference</td>
</tr>
<tr>
<td>Work hours</td>
<td>More</td>
<td>More</td>
<td>—</td>
</tr>
<tr>
<td>Teacher absences/school closings</td>
<td>Less</td>
<td>Less</td>
<td>—</td>
</tr>
<tr>
<td>Use of alternative pedagogy</td>
<td>—</td>
<td>Less</td>
<td>—</td>
</tr>
<tr>
<td>Salary</td>
<td>Less</td>
<td>Less</td>
<td>—</td>
</tr>
<tr>
<td>Receives incentives</td>
<td>—</td>
<td>—</td>
<td>More</td>
</tr>
</tbody>
</table>

— is no data are available.

Source: Vegas and Umansky 2005.

It is possible that the lower wages and more difficult working conditions of teachers in EDUCO and PROHECO schools discourage talented teachers from choosing to work in these schools. In El Salvador and Honduras, teachers in community-managed schools have fewer years of experience than those in traditional schools and receive lower salaries. In Honduras, PROHECO teachers have fewer years of schooling and report facing significantly more problems in receiving their salary on time. In Nicaragua, by contrast, where autonomous schools were pre-existing and are predominantly urban, teachers in autonomous and traditional schools have roughly the same education and experience level.

There are also differences between locally managed and traditional schools in teacher reports of how many hours they work at the school. In El Salvador, teachers in EDUCO schools report working more hours than their colleagues in traditional schools. Some of the additional hours may be due to statistically significant increased time spent meeting with parents. In Honduras, there is limited evidence that PROHECO teachers spend fewer hours on administrative tasks and more hours teaching. PROHECO teachers also assign more homework and have smaller classes than teachers in similar traditional schools, two factors that are often correlated with higher achievement. These examples lend credence to the idea of greater efficiency and teacher effort in decentralized schools. In Nicaragua, there are fewer statistically significant differences between treatment and control schools than in Honduras and El Salvador, although this may be partially due to data constraints.

While most of the first-order effects would be expected to boost teacher quality and student learning, in Honduras no evidence was found that teachers in community-managed schools are more motivated than are teachers in traditional schools. PROHECO teachers are not significantly more likely to feel valued, respected, or safe. The only significant difference between teachers in
PROHECO schools and teachers in traditional schools on questions dealing with teacher motivation is that traditional school teachers are more likely to feel close to parents. The decreased closeness of teachers and parents in PROHECO schools may result from the new power of parents on school councils to fire teachers.

In classroom processes, teacher planning, and school environment there are very few differences between teachers in PROHECO schools and those in traditional schools. One important exception, according to student surveys, is that teachers in PROHECO schools are significantly more likely to use traditional or frontal teaching methods. PROHECO students report that their teachers are more likely to use dictation and teach from the blackboard and are less likely to have students work in groups. This finding suggests that PROHECO schools may have been less successful than traditional schools at implementing Honduras' national pedagogical reform for more dynamic pedagogy.

With this understanding of some of the first-order effects of the Central American school-based management reforms on school management and teacher characteristics and behavior, the following section turns to how the reforms and their impact on teachers are associated with student learning outcomes.

Implications of First-order Effects of School-based Management on Student Learning

Several studies have shown that school-based management can improve student learning, while others have shown that it can increase inequality and harm learning, particularly for the poor. But few studies have analyzed the mechanisms that explain such mixed results. Do the greater roles of parents or principals improve accountability and communication and result in better, more relevant teaching? Do longer teacher working hours, better training, or different teaching methods affect learning? Where a school-based management reform fails to improve student learning or actually harms student learning, do any of the changes brought about by the reform explain this failure?

The three Central American reforms indicate that school-based management reforms can have varied impacts on student learning. Effects on management and teachers account for many of the mixed effects on learning (table 3).

In Honduras, PROHECO students score significantly higher on math, science, and Spanish exams than do students in traditional schools. The benefits of PROHECO are partly explained by the qualities and characteristics that differentiate PROHECO schools from otherwise comparable poor rural non-PROHECO schools. Specifically, the more hours per week a teacher works—higher in PROHECO
Schools—the greater student achievement in all three subjects. The frequency of homework—also greater in PROHECO schools—is associated with higher achievement in Spanish and math. Finally, smaller classes and fewer school closings—as in PROHECO schools—are related to higher student achievement in science.

In El Salvador, EDUCO students also perform better in Spanish than do students in traditional schools, background factors being held constant. Again, the benefit of EDUCO appears to come partly from the observed differences between teachers in EDUCO schools and teachers in traditional schools. Specifically, the greater amount of time EDUCO teachers spend meeting with parents partially explains why EDUCO students outperform their counterparts. There is some evidence that EDUCO students may also outperform traditional school students in math and may be less likely to be absent, but these results are much more tentative.

In Nicaragua, the School Autonomy reform does not appear to have led to increased student learning. Although students in autonomous schools have significantly higher average test scores in mathematics in third grade, by sixth grade autonomous school students score lower than do students in traditional schools in both Spanish and mathematics once background characteristics are taken into account. Furthermore, none of the differences in test scores between schools appear to result from changes in school management or teaching due to the reform, with one exception: greater technical assistance to teachers from principals appears to improve third-grade students' scores in Spanish. Why sixth-grade students perform worse than their counterparts in traditional schools is not explained by any of the variables included in the Nicaraguan analysis.
Policy Implications

The three school-based management reforms in Central America have had varying degrees of success in improving student learning. The PROHECO reform in Honduras seems to have had the most success in improving students’ test scores in multiple subject areas. It did so through the instrumental means of keeping schools open more, giving more homework, having smaller classes, and having teachers work more hours per week. By the sixth grade, however, students in Nicaragua’s autonomous schools are performing worse than those in traditional schools. What explains these differences, and what lessons can be drawn from school-based management reforms in Central America?

1. It cannot be taken for granted that school-based management will result in well-run schools and empowered communities. If school-based management is offered as a solution to pressing educational problems, school councils must be empowered to make meaningful improvements in schools and must be capable of doing so. For example, the evidence from El Salvador suggests that in many important areas EDUCO schools may have less decisionmaking power in practice than is intended.

Clearly, the effectiveness of school-based management will depend to a large extent on the capacity of local stakeholders to manage schools. The ability of communities and parents to identify, hire, and retain good teachers and promote good teaching practices can vary widely. They are not necessarily any more effective—perhaps significantly less so—than competent centralized management mechanisms. For example, PROHECO teachers in Honduras are significantly more likely to report month-long delays in receiving payments than are teachers in traditional schools. This suggests that the bureaucracy and inefficiencies of some centralized education systems may be replaced by general mismanagement of resources in decentralized systems.

To avoid these pitfalls, school-based management reforms may want to avoid placing large amounts of power in the hands of any one stakeholder. The excess of power granted to school principals in Nicaragua, analyzed elsewhere, may be problematic (King and others 1996). It is also critical for effectiveness and equity that all school councils be ready and able to perform their duties and functions. This requires both building capacity and providing resources, especially in poorer communities and communities with less management experience and capacity.

2. Teacher behaviors appear to respond to the incentives created in locally managed schools. There is evidence that, as a result of school-based management, teachers are working more hours, assigning more homework, and meeting more often with parents. These are promising changes, many of which appear to
contribute to increased student learning. In contexts of low teaching quality, these changes are critical steps in a positive direction.

(3) School-based management appears not to be improving teaching methods or teacher professionalization. The reforms have clearly created new incentives for teachers, but most of these changes are capacity-utilization changes—smaller classes, more hours, fewer closings—not changes in the kind of teaching that is taking place inside classrooms. The reforms do not appear to result in positive changes in teaching methods, curricula, school environment, or teacher motivation level. In Honduras, PROHECO teachers appear to be more dependent on antiquated teaching methods and less likely to feel respected and safe at work than teachers in traditional schools. Teachers in locally managed schools in Nicaragua and Honduras report having less authority than do teachers in traditional schools. Complementary reforms or modifications of school-based management reforms are needed that result in greater teacher empowerment, better teacher skill development, and improved teaching methods.

(4) There may also be undesirable incentives in community-managed schools that could have negative implications for teachers. Lower salaries as in Honduras, and less secure jobs in all three countries may discourage talented teachers from working in locally managed schools. Smaller class sizes in Honduras and technical assistance from principals in Nicaragua may in part be compensating for teacher weaknesses.

(5) The policy context surrounding the reform is bound to affect its impact. In Nicaragua, the School Autonomy reform was implemented along with monthly student fees for basic and secondary education. These changes were perceived by some of the general public, however inaccurately, as an effort to privatize basic education in the post-Sandinista era and may have limited public acceptance of the reform. In all three countries, teachers’ unions have tended to resist the school-based management reforms, which are seen as a threat to union power and teachers’ job stability. The negative publicity and resistance from these unions may also limit the effectiveness of the reforms.

Conclusions

Previous research on educational decentralization has explored the impact of various decentralization reforms on indicators such as student learning and completion. But these analyses have largely looked at reforms as a “black box” rather than investigated the chain of effects that decentralization reforms put into action. The analysis of school-based management reforms in Central America reported here differs from previous work in that it first looks at how the reforms affected school management and teacher characteristics and behavior and then evaluates
how these changes in management and teachers are associated with improvements or deterioration in student learning.

Education reforms, even those not specifically designed to affect teachers, influence the characteristics of teachers and their daily work in classrooms. The analysis of school-based management reforms in Central America summarized in this article indicate that these reforms can have effects on teacher salaries, feelings of authority and autonomy, and hiring and firing practices, all of which generate strong teacher incentives. For example, the authority of EDUCO school councils to hire and fire teachers was found to be an important factor in better outcomes for EDUCO students compared with students in traditional schools serving similar populations in El Salvador. The positive impact of education reforms can be strengthened by understanding of teacher incentives and thinking through the implications of diverse kinds of reforms for teachers.

Note
Ilana Umansky is assistant director for International Education and Research at Sesame Workshop; her e-mail address is ilana.umansky@sesameworshop.org. Emiliana Vegas (corresponding author) is senior education economist in the Latin American and Caribbean Human Development Department at the World Bank; her e-mail address is evegas@worldbank.org. The authors are grateful to Ariel Fiszbein, Beth King, Peter Moock, Guillermo Perry, Jeff Puryear, Halsey Rogers, Jaime Saavedra, Eduardo Velez Bustillo, and three anonymous reviewers for helpful comments.

References


Incentives to Improve Teaching: Lessons from Latin America. World Bank, Washington, D.C.


Using Global Positioning Systems in Household Surveys for Better Economics and Better Policy

John Gibson • David McKenzie

Distance and location are the important determinants of many choices that economists study. Economists often rely on information about these variables that is self-reported by respondents in surveys, although information can sometimes be obtained from secondary sources. Self-reports are typically used for information on distance from households or community centers to roads, markets, schools, clinics, and other public services. There is growing evidence that self-reported distance is measured with error and that these errors are correlated with outcomes of interest. In contrast to self-reports, global positioning systems (GPS) can determine location within 15 m in most cases. The falling cost of GPS receivers makes it increasingly feasible for field surveys to use GPS to more accurately measure location and distance. This article reviews four ways that GPS can lead to better economics and better policy by clarifying policy externalities and spillovers, by improving the understanding of access to services, by improving the collection of household survey data, and by providing data for econometric modeling of the causal impact of policies. Several pitfalls and unresolved problems with using GPS in household surveys are also discussed. JEL codes: C81, O12, R20

Distance and location are important determinants of many choices that economists study. For example, in the von Thünen model, distance to market determines landowners' decisions about what crop is most profitable to produce. Studies of child labor market activity find distance from urban areas to be an important determinant of both schooling and work decisions (Falchamps and Wahba 2006). In migration models, greater distance between origin and destination implies larger migration costs and reduced migration flows (Borjas 2001).
Although information on location and distance can sometimes be obtained from secondary sources, economists often rely on information that is self-reported by respondents in surveys. These self-reports are typically used for data on the distance from households or community centers to roads, markets, schools, clinics, and other public services. Evidence is presented here showing that self-reported distances and areas are measured with considerable error and are often correlated with outcomes of interest. In contrast, global positioning systems (GPS) can determine location within 15 m most of the time.

GPS locations are determined from satellites (currently 30) with precise atomic clocks. The satellites orbit about 20,000 km above the earth and send unique radio signals with a time-stamp. A GPS receiver uses the time delay between transmission and reception to calculate the distance to each satellite and to calculate the latitude and longitude of the location by triangulation. More precise calculations, including elevation, can be made if four satellites are in view (El-Rabbany 2006). Accuracy depends partly on the GPS receiver's unobscured view of the sky and partly on the quality of the receiver used to process the satellite signal. Consumer-grade GPS receivers are accurate to within 15 m 95 percent of the time, with further improvements in accuracy to about 3 m achievable by using differential GPS, which augments satellite information with information from a local reference station.

Two principal factors have dramatically increased the feasibility and usefulness of collecting GPS information in household surveys. First, on May 1, 2000, the U.S. military turned off selective availability, which had introduced random errors of up to 100 m in the civilian signal. The removal of selective availability allowed more accurate measurement, increasing the range of possible applications. Second, the cost of a basic GPS receiver has fallen to under $100, bringing it within the budget of most household surveys. Coverage and precision will improve even further with the launch of the European GALILEO system, expected to be operational by 2008.

The Demographic and Health Surveys (since 1997) and the Indonesia Family Life Survey are the two surveys that are well-known to economists who have used GPS to geo-reference the locations of community centers (and hence cluster of households, given the sample design). GPS has been used to locate individual households (and enterprises) in a few recent World Bank surveys, including the Rural Investment Climate Surveys in Indonesia and Sri Lanka and the Living Standard Measurement Study surveys in Albania and Tanzania. Still, the majority of household surveys in developing economies do not geo-reference communities or households, in part because of a lack of information about the benefits.

An alternative to using GPS is to use secondary sources of data on locations. In developed countries, postal addresses are widely used. For example, the United Kingdom has 2.1 million post codes for 26 million addresses. Post codes are a
very accurate proxy for household location since it is possible to get map-grid references to the nearest 100 m for most post codes. Duranton and Overman (2005) use these very detailed location data to examine the location patterns of manufacturing. However, few developing economies have detailed post codes. In developing economies, face-to-face interviewing predominates (telephone interviewing is the norm in developed countries), making it quite feasible for field teams to gather GPS data as a part of their usual survey workload.

Another source of secondary data is remote sensing, which gathers data from a sensor mounted on an aircraft or satellite. These data are typically used in studies of land cover, dealing with topics such as deforestation (Deininger and Minten 2002) and urban sprawl (Burchfield et al., 2006). The unit of analysis is the pixel or picture element, which determines the size of the smallest landscape feature that can be distinguished and mapped. Typical sizes are $30 \times 30$ m or $1 \times 1$ km grids. However, these grid cells are not individual agents, and using data at this level may involve aggregating across decision-makers and may result in the ecological fallacy of drawing inferences about the behavior of individuals from analyses based on grouped or area-level data (Freedman 2001). A better matching of the spatial scale of the decision process and the scale at which measurement is carried out (Anselin 2002) may come from surveying individual decision-making agents and using GPS to link them to other spatial data.

Geographic information systems (GIS) enable such linking of different layers of data. As GIS can be considered simply a tool for combining, manipulating, and displaying spatial information captured in a variety of ways, including through GPS, a broader view sees it as a part of an emerging geographic information science (Goodchild 1992) that may enable researchers to discover new relationships for geographically referenced information. Some of the literature reviewed here relies more heavily on GIS than on GPS but still shows the types of analyses that could be facilitated by the more frequent use of GPS in household surveys.

This article reviews four ways that GPS can lead to better economics and better policy by clarifying policy externalities and spillovers, by improving the understanding of access to services, by improving the collection of household survey data, and by providing data for econometric modeling of the causal impact of policies. The article also discusses some pitfalls, unresolved problems, and ongoing research.

**Four Ways Using GPS can Lead to Better Economics and Better Policy**

The use of GPS can lead to better economics and better policy in at least four ways.

John Gibson and David McKenzie
Using GPS can Clarify Policy Externalities and Spillovers

The spatial proximity of one household to another may be of direct interest, particularly for understanding interactions between households, the role of social networks, and the potential spillovers from policies that treat some households and not others.

One example of interactions between households that researchers might want to study is the possibility that households learn from each other’s actions. Conley and Udry (2005) study learning in the context of the decision to cultivate pineapple in Ghana and to determine how much fertilizer to apply. They note that the classic identification problem is that a farmer’s greater likelihood of adopting a new technology soon after neighbors have done so might be a consequence of some unobserved variable that is spatially correlated, such as soil type, pests, or topographic features, rather than a result of genuine learning. They use GPS to define the geographic neighbors of a given plot as those within 1 km of the center of the plot, and they collect data on whom farmers talk to (informational neighbors). Controlling for the deviation of a farmer’s input from that of his geographic neighbors, Conley and Udry could then identify learning through the impact of informational neighbors’ choices. They also find evidence of positive spatial correlation in unobserved shocks to the productivity of fertilizer, highlighting the importance of controlling for geographic effects when examining learning.

McKenzie, Gibson, and Stillman (2007) also use GPS to study learning from neighbors, examining how emigrants’ negative employment experiences affect the expectations of would-be emigrants. The would-be emigrants were all unsuccessful in a random ballot in Tonga that offers an opportunity for ballot winners who obtain employment to move to New Zealand. When subsequently interviewed about their employment and income expectations had they moved to New Zealand, the would-be emigrants greatly understated employment rates and incomes compared with actual outcomes for emigrants. One explanatory factor is that many ballot winners who did move to New Zealand found that their initial job offers were no longer available, and news of this negative outcome appears to have flowed back to the would-be emigrants in Tonga. Specifically, if all ballot-winning emigrants within a 6 km circle (based on GPS measurements) of the ballot losers did not take up their initial job offers in New Zealand, the employment expectations of the ballot losers were 19.6 percentage points lower.

The standard approach for evaluating the impact of a policy is to compare outcomes for those subject to the policy with outcomes for a comparable group not subject to that policy. However, as Miguel and Kremer (2004) point out, this can give misleading results when there are externalities. They investigate the impact of a deworming treatment in schools in Kenya. Using GPS distances at the school level, they control for the number of primary school pupils within a certain
distance of the school and use the number of treated pupils within this distance to measure health spillovers. They find that naïve estimates that fail to take externalities into account would underestimate the program treatment effects, leading to the mistaken conclusion that deworming is not cost-effective. Such an approach could be extended by obtaining GPS locations of the residence of each child, which could then be used to construct a child-specific measure of exposure to treated and nontreated children. This would provide more variation in the extent of spillover, which could be used to examine the heterogeneity in treatment effects. A recent example of research examining spillovers at the individual level is provided by de Mel, McKenzie, and Woodruff (2007). They conduct a randomized experiment in which grants of capital stock are given to a randomly selected sample of Sri Lankan firms. They then estimate the impact of this treatment, controlling for the number of treated firms within 500 m and 1 and 5 km of each surveyed firm.

Information on the Spatial Distribution of Population and Services is Essential to Improving Understanding of Access to Services

One of the most common uses of GPS information in developing economies has been to measure access to infrastructure and social services, particularly health care. For example, Perry and Gessler (2000) use GPS to measure access from communities to primary health-care facilities in Andean Bolivia and use the results to propose an alternative model of health distribution in the study area.

In addition to providing purely descriptive measures of access, GPS data on distance and travel time can identify barriers to the use of services. In examining the influence of accessibility to family planning on choice of contraceptive device, Entwisle et al. (1997) demonstrate two advantages of GPS over survey-based measures of access. First, surveys often collect data on family planning accessibility only for certain political or administrative boundaries, establishing, for example, whether there is a facility in the village. However, facilities in neighboring administrative units may be closer. Using geo-referenced data allows more flexible specification of boundaries unconstrained by administrative definitions. Second, respondents in surveys often report travel times to health facilities in 30-minute increments, whereas GIS gives no time clumping, allowing better specification of functional form.

Gibson et al. (forthcoming) examine the use of different channels for receiving remittances in Tonga. Transaction costs on money transfers are much higher for Western Union than for withdrawals from an automated teller machine (ATM). There are eight ATMs on the main island of Tongatapu and five Western Union branches, so a simple measure of branches and ATMs per capita would suggest that ATMs are more accessible. However, when GPS coordinates of the ATMs and
Western Union branches are combined with village-level population information from the census and a digitalized map of the road network to measure the share of the population within different travel distances of the two financial channels. The Western Union branches are shown to be more dispersed and to offer better access. The branches cover 97 percent of the population within a 10 km travel distance, whereas the ATMs cover only 77 percent of the population within this distance (figure 1). This combination of GPS data collection and mapping software can be particularly effective in illustrating access in a form that policymakers can readily use, as shown in figure 1.

Recent health applications combine distance with measures of health infrastructure quality. Hong, Montana, and Mishra (2006) use the 2003 Demographic and Health Surveys in Egypt to look at the relationship between the use of IUD contraceptives and the quality of family planning services. They link each household to the nearest family planning clinic within 10 km and then use detailed survey data to measure the quality of the facility. Rosero-Bixby (2004) uses GPS data on census tracts and locations of health facilities in Costa Rica to assess improvements in access following health reforms, measuring access through a combination of distance and services provided by the facility. He notes that households may not necessarily use the nearest facility, particularly if it is of low quality, and that using GIS enables calculating such measures as the density of services that meet a standard quality within a specified radius.

A limitation of these health studies is that they measure distance only at the community level, whereas households on opposite sides of a village or town may each be closer to a different facility, either of which may be in another community. A second limitation is that distance to health facilities could be correlated with a host of other unmeasured factors, such as poverty, disease environment, and other infrastructure, which could also affect health decisions.

Using GPS Can Improve the Collection of Household Survey Data

GPS is also being used to improve the quality and cost-effectiveness of household survey data. GPS is being used in several phases of data collection, from the development of a sample frame to quality control and follow-up surveys. More accurate and cost-effective surveying enables researchers to carry out better analysis and provide better evidence-based advice to policy-makers.

Household surveys require an accurate sample frame. The most common approach involves using a recent census to select enumeration areas. However, censuses may become outdated during periods of rapid urbanization and are of little use in drawing samples in post-conflict countries that have not had a census for decades. Afghanistan plans to complete a census in 2007, its first since 1979, and Lalasz (2006) reports that 15 countries have not taken a census.
Figure 1. Service Areas for ATMs and Western Union Branches for Tongatapu, Tonga

Source: Gibson et al., (forthcoming), figure 4.
since 1990. The traditional solution is to do area sampling, in which enumerators list all households in a well-defined block, such as a village or an urban area bounded by certain streets. These blocks are determined largely by convenience in defining and locating them and can be expensive to enumerate.

Landry and Shen (2005) show how GPS can be used to do area-based sampling quickly and cheaply, since enumeration areas can be defined by spatial coordinates and made arbitrarily small. Landry and Shen consider the problem of surveying in China, where household registration lists are widely used as sample frames. Widespread migration from rural areas means that many households are unlikely to be found on these registration lists. They use GPS to survey randomly chosen 54 × 54 m² (approximately one square second) spatial blocks and find that 45 percent of the households reached were not on household registration lists.

A potential problem with this approach is that the sample size is not known until after data collection, since the number of households within a spatial block will not be known in advance. Landry and Shen use the existing population data to create a rough population model of Beijing. Since the number of dwellings within their spatial units was four times as large as they had budgeted for, they administered their questionnaire to just a quarter of the units.

Aerial photography is likely to alleviate such problems in the future. For example, Cowen and Jensen (1998) extract information on individual dwelling units in a 32-block census area in South Carolina from aircraft multispectral data. They find that the dwelling unit data derived from remote sensing had a correlation of 0.91 with data derived from the census. As satellite imagery continues to improve in resolution and fall in price, it appears likely that the combination of remote sensing and spatial sampling will become the standard for constructing sample frames in situations where reliable census or registrar data are not available.

Kumar (2007) describes combining remote sensing and GPS for drawing samples in a survey of 1,600 households spread across different air pollution zones in Delhi and India. He partitions the study area into strata according to air pollution levels (obtained by remote sensing) and proximity to main point sources of air pollution. Random points are then simulated using GIS techniques (weighting by size of residential area in each strata), and GPS is used to navigate to the households located at each selected point, and the households are asked to participate in the survey. This method of creating a frame and drawing a sample should be more efficient than simply imposing a regular grid on the study area, since air pollution is irregularly distributed over space.

Visualization of the locations at which sampling has occurred can provide a useful form of quality control to ensure that interviewers conduct surveys where they are supposed to and to check whether any dwellings are inadvertently missed. In 2004, Timor-Leste became the first country to use GPS units in its
census to record the locations of all households. USAID/Timor-Leste (2004) reports that survey managers checked the GPS points visited by the census teams against detailed aerial photographs to detect any areas missed in the enumeration. Census undercounting matters for a variety of policy purposes, including the allocation of federal money and political representation. Undercounting can be particularly high in developing economies. Lalasz (2006) reports that the 1991 Census is thought to have undercounted Nigeria's population (officially put at 89 million) by perhaps 20 million people. The use of GPS can show where undercounting has occurred and help survey managers to reduce it.

GPS can also reduce the cost and time to relocate dwellings for follow-up surveys. Follow-ups may be needed to allow field managers to check for errors by enumerators and to collect panel data. In many developing economies, the lack of street addresses, especially in densely populated urban areas, and changes in administrative boundaries between waves of a survey can make identifying the same dwelling or household time-consuming. A pilot study by Dwolatsky et al. (2006) tracing patients who left a tuberculosis control program in South Africa shows the potential for using GPS to relocate dwellings. They find that it takes 20–50 percent less time to locate a home using a customized personal digital assistant linked to GPS than using residential addresses. The main limitation is that this was a small pilot study of only 20 houses, so further experiments are needed to confirm the promising results.

When panel surveys attempt the more difficult task of tracking individuals rather than dwellings, GPS can be very useful for tracking people who had previously been co-residents. For example, the 2004 Kagera Health and Development Survey in Tanzania used GPS to record the locations of 2,700 households containing members who had been in the baseline sample of 900 households first interviewed in 1991–1994 (Beegle, De Weerdt, and Dercon 2006). Measures such as how far people have moved from their baseline village center or from households with members who had been co-residents in the baseline surveys can be related to various socioeconomic characteristics.

Finally, collecting GPS data for households enables linking the household data set to other surveys and data sets. There is considerable option value in doing this, since many potential uses of the data will not be known at the time the survey data are collected.

**GPS Can Be Used to Provide Data for Econometric Modeling of the Causal Impacts of Policies**

Most empirical work in development economics aims to identify the effect of a particular variable of interest, X, on a particular outcome, Y. A standard concern is that there are other variables that are correlated with X and that also affect Y.
Failure to control for these variables gives biased results. One of the most basic uses of GPS is to allow researchers to better control for geographic and locational characteristics in their regressions. Such characteristics are increasingly found to be relevant to outcomes of interest for development economists and practitioners. For example, Deininger and Minten (2002) obtain data from a GIS on soil quality, rainfall, elevation, slope, and other geographic features and find that higher levels of poverty are statistically associated with greater likelihoods of deforestation. However, when they re-estimate the model without GIS data, they find poverty to be associated with lower levels of deforestation. The problem is that poor people live on worse-quality land, which limits the benefits (such as agricultural income) from the deforested land, so controlling and not controlling for land quality give opposite results.

Propensity-score matching has become a popular tool for investigating policy impacts (see Ravallion, forthcoming, for a recent review). The idea is to compare individuals subject to a policy with similar individuals not subject to it. Typical variables used for matching are household socioeconomic characteristics and an often crude set of community-level variables. Brady and Hui (2006) argue that GIS can be used to more explicitly include geography in matching. They present three arguments for doing so: much individual data that would be useful for matching is unmeasured and place can serve as a proxy for unmeasured individual characteristics; nearby places are more likely to share community characteristics, such as culture, trust, and government ability; and geographic matching can be visually persuasive when sudden changes in outcomes occur across administrative borders and a program is in one community and not in its neighbor.

Nevertheless, Brady and Hui acknowledge that in some cases the most comparable places in terms of cultural or socioeconomic characteristics may not be geographically close. Therefore, matching must be based on more than geography. Although the literature on the U.S. labor market emphasizes the importance of comparing participants in training programs with nonparticipants from the same local labor markets (Heckman, Ichimura, and Todd 1997), the literature has generally not explicitly included geographic proximity as a criterion when matching individuals in different communities. As more surveys include GPS coordinates, this will become increasingly possible.

The two examples above highlight the ability of GPS to help researchers better control for (potentially) observable characteristics. More controversial is the use of distance or other geographic variables as instruments in instrumental variables estimation. Oster (2006) uses distance to the Democratic Republic of Congo as an instrument for HIV prevalence when examining the response of sexual behavior to HIV prevalence rates in Africa. McKenzie, Gibson, and Stillman (2006) use GPS-measured distance from a household in Tonga to the New Zealand immigration office in Tonga where application forms must be deposited as an
instrument for migration, when looking at the effect on income of migration to New Zealand. Olken (2006) uses GIS data on community locations and geography to study the impact of television and radio on social capital in Indonesian villages. He argues that geography leads to differences in signal strength in different villages because of mountains located between some villages and the transmission towers, but after controlling for other variables, he concludes that this geography has no independent effect on social capital.

However, using distance as an instrument is subject to potential problems. One is that distance to borders and major cities is also likely to determine access to markets, schools, health facilities, and other infrastructure, which can influence economic behavior. Also, people, villages, and cities are not randomly allocated in space. As a result, distances usually incorporate the results of behavioral choices, some of which may affect outcomes. The standard response to such concerns is to include as many other geographic controls as possible. For example, Olken (2006) controls for elevation, district fixed effects, and distance and travel time to major cities and uses a physical model of radio transmission that predicts how signal strength should vary with topography. But as with all instrumental variables, even after including such controls, a case needs to be made for why the exclusion restriction should hold—why should one believe that unobserved geographic features are not also influencing the outcomes of interest?

A second potential concern with the use of distance as an instrument arises when the response of interest varies across individuals. Even if the exclusion restriction holds, the instrumental variables estimator will identify only the local average treatment effect (LATE) in this case. As Heckman (1997, p. 451) notes about the use of distance to the nearest school as an instrument for schooling, "LATE estimates the effect of variation in distance on the earnings gain of persons who are induced to change their schooling status as a consequence of commuting costs that vary within a specified range." Whether estimation of such a parameter interests policy-makers is a matter of some doubt.

There is less concern about this issue when most individuals respond to distance in a similar manner. McKenzie, Gibson, and Stillman (2006) find that 98 percent of individuals who did not apply for the migration lottery in Tonga gave lack of information as the main reason. Living closer to the consulate office results in better information for most individuals; so distance might be expected to change migration status for most individuals in the sample. Indeed, they find that using distance as an instrument gives an estimated income gain from migration of within 2 percent of that obtained from the experimental estimator provided by a migration lottery. Thus, this is an example where using distance as an instrument provides reliable results.

Similarly, it may be that shocks to local environments, as captured by remote sensing data in two time periods, provide a more defensible identification
strategy. For example, households can be linked to areas of flooding, earthquakes, tsunamis, and other such shocks. One practical constraint is that converting the satellite images to usable data is costly and time-consuming with the current manual techniques.

A final use of GPS is in spatial econometric models. Many unobserved variables, such as climate and soil in agricultural settings, are spatially correlated, leading to spatial autocorrelation in the error term of regression equations. Failure to account for this structure in the error terms will lead to the use of incorrect standard errors for inference, possibly leading to the conclusion that a policy has a significant effect when it does not, or vice versa. Distances between observations obtained through GPS can be used to account for spatial autocorrelation in the error term of the regression equation. Case (1991) and Conley (1999) provide procedures for doing this.

How Much Improvement Does GPS Give Over Self-Reports, and Is a Straight Line Good Enough?

A natural question that arises when distances from households and communities to other households, communities, or infrastructure need to be known is whether GPS should be used to measure these or whether self-reports in household surveys are accurate enough. A follow-up question is whether a simple straight-line distance (as the crow flies) is sufficient or whether the GPS coordinates should be integrated with GIS information on transport routes and topography to measure travel distances and travel times.

The consequences of mismeasuring distance depend on how distance is going to be used and on how badly it is mismeasured. If measurement errors are classical (independent across individuals and over time and uncorrelated with individual characteristics), then when distance is used as a regressor, as in studies of access, the effect will be to understate the impact of distance (attenuation bias). Using distance as an instrument with classical measurement error will lower the power of the instrument, potentially giving rise to weak-instrument concerns, but it will still result in consistent estimates.

However, there are strong reasons to believe that measurement errors are not random. Entwisle et al. (1984) note as an example that if people are asked to report travel times to a health provider, those who currently use that health resource will have more accurate knowledge than those who do not. Thus, the measurement error is likely to be correlated with usage patterns, a problem for investigating the impact of distance on use. Indeed, Andrabi et al. (2007) report in their survey in Punjab and Pakistan that many households do not even know
the name of the nearest school, let alone its location. If the measurement error is correlated with socioeconomic variables that also affect the outcome of interest, then the mis-measured distance will also give inconsistent instrumental variable estimates.

Few studies systematically compare self-reports of distance and travel times to GPS measurements, particularly in developing economies. For this study, a recent World Bank survey of owners of microenterprises and small enterprises in Bolivia was used to make the first known comparison of self-reports of physical distance to GPS-measured straight-line distances. Firm owners are required to register at the local branch of the national tax system, but only 30 percent of the firms in the sample had registered. Firm owners were asked the distance in kilometers to the nearest tax office, which in heavily urban areas could be compared with the straight-line distance taken from GPS coordinates of the firm and tax office. More than half the firms said that they did not know the distance, with lack of a response strongly correlated with whether a firm had registered: 68 percent of unregistered firms answered that they did not know, whereas only 25 percent of registered firms did not know.

Figure 2 shows a scatterplot between the reported and measured distances for firms within 10 km of the tax office, conditional on firms also reporting the distance as 10 km or less. Pearson’s correlation between reported and actual distance is still only 0.39, and the Spearman’s is 0.31. The degree of measurement error conditional on giving a self-report is not significantly related to whether the
firm is registered or to the age, gender, marital status, or education of the firm owner. However, men, more educated individuals, and registered firm owners are more likely to report a distance.

Self-reports of time can also contain systematic measurement error. Escobal and Laszlo (2005) compare the self-reported time for agricultural producers in Peru to get to the nearest population center with the true travel time. The true time is measured by having surveyors walk with a random sample of respondents and time their journeys, following the same route and pace as the respondent and using GPS to measure latitude, longitude, altitude, and distance. GIS is then used to account for terrain and compute travel time for those in the survey who were not accompanied by the surveyor. They find that respondents consistently underreport the time to reach the center. For example, among coffee producers in the Selva, mean self-reported time was 6.7 minutes, compared with a mean true time of 13.0 minutes. The correlation between self-reported time and true travel time is only 0.28 for coffee producers, 0.29 for potato farmers and −0.08 for rice farmers in their sample. Furthermore, Escobal and Laszlo find that measurement errors are correlated with socioeconomic variables. Not surprisingly, individuals who own a watch give more accurate reports of travel times. They also find a negative correlation between measurement error and education so that more educated people have less measurement error.

These results strongly suggest that self-reported distances will be misleading, with measurement errors correlated with outcomes of interest. GPS coordinates can be used to give more accurate measurement. The simplest approach is to calculate the straight-line distance between points. This has the advantage of computational simplicity and does not require additional geo-coded information on transport networks or topography. Alternatively, users can combine GPS point coordinates with information on the location of transportation routes and perhaps with information on road quality and topography to measure exact travel distances and predict travel times.

The correlation is likely to be much higher between GPS straight-line travel distances and exact travel distances than between self-reported distance and GPS distances. For example, McKenzie, Gibson, and Stillman (2006) compute the distance in meters from each household in their sample to the New Zealand immigration office in Nuku‘alofa. The Pearson correlation between the straight-line and exact travel distance based on road networks is 0.82, and the Spearman correlation is 0.78. They find the absolute percentage measurement error to be correlated with whether an individual migrates and with income from work in Tonga. The measurement error is greater for individuals located in more remote areas (on the other side of a lagoon), and this remoteness in turn is correlated with economic behavior. Nevertheless, the size of this error is small enough in application that there is no difference in the instrument variable estimates
Between straight-line distances and road distances. The income gains from migration are estimated to be $280 (standard error $122) using the straight-line distance as an instrument for migration and $281 (standard error $101) using the road distance.

More generally, the difference between straight-line and road distance measures will be larger when geographic features such as mountains, lakes, lagoons, and rivers lie between a household or village and the location of interest. Thus, the difference between straight-line and road distances will be correlated with the remoteness of a location, which in turn is likely to affect many variables of interest. As a result, road distances are preferred to straight-line distances where possible.

Furthermore, since a curve between two points is always longer than a line, travel distances will be longer than straight-line distances. As a result, measures of access based on straight-line distances will overestimate the proportion of the population that is covered by a given service. This is demonstrated by Noor et al. (2006), who examine health coverage in Kenya, where the government has set a target of ensuring that no one lives more than one hour away from effective health services by 2010. Estimates based on straight-line distances—the standard approach to coverage—indicate that 82 percent of the population is within one hour of government health services. Adjusting for the travel network drops coverage rates to 63–68 percent. This would mean that 19 million people rather than 25 million are currently covered.

Pitfalls and Unresolved Problems

Several problems need to be avoided or overcome in using GPS in household surveys.

Interviewer Error

Although taking GPS readings is straightforward, it requires good training to avoid creating another source of measurement error. One method of improving accuracy in readings is to have interviewers take multiple readings for the same location and then use their average. Some GPS receivers have a built-in function for doing this. The guidelines for collecting GPS data in Demographic and Health Surveys recommend taking multiple readings within a five-minute period and averaging them (Montana and Spencer 2004). The fieldguides of Spencer et al. (2003) and Montana and Spencer (2004) provide a good starting point for researchers planning to use GPS in household surveys. The guides recommend at least 60 minutes of outdoor hands-on training with the GPS units.
Datum and Coordinate Projection Problems

A spheroid approximation of the shape of the earth is used to solve geodetic problems for point location, and this surface, its origin, and the orientation of its latitude and longitude lines make up a "geodetic datum." GPS receivers typically use the World Geodetic System 1984 (WGS84) datum, which is a geocentric datum (its origin coincides with the center of the Earth) designed for making worldwide measurements. However, there are hundreds of other datums that may use a different center, spheroid, or reference point on the Earth's surface in order to be locally more accurate. Coordinate values resulting from interpreting latitude, longitude, and height values based on one datum as though they were based on another can cause position errors of up to 1 km (Ramachandran 2000), although the discrepancy will typically be less. Bennett (2006) gives the example of walking around Tiananmen Square in Beijing with a GPS receiver and then importing the measurements into Google Earth, which shows a path offset by approximately 14 m because of the difference between WGS84 datum and the datum used by Google Earth.

Another common source of error comes from mixing geographic and projected coordinate systems. Projected coordinates overcome the problem that latitude and longitude are not constant units, so Cartesian geometry cannot be used to measure either distances or areas when work is being done with latitude and longitude coordinates. Projected coordinate systems convert latitude and longitude coordinates from the Earth's three-dimensional surface onto a two-dimensional map. Consequently, if location data from a GPS (which are for a three-dimensional surface, and so use a geographic coordinate system) are combined with two-dimensional map data (in a projected coordinate system), lining up the various data layers requires conversion. This is easily done in a GIS. Building up different layers of data for households, villages, and features of interest such as roads, coastlines, rivers, and public services adds value to the information in each layer. But it is surprisingly easy to have unmatched coordinate systems, because metadata, which should tell users about the coordinate system and datum used, are not always included with existing geographic data. For example, one of the authors digitized a road around the edge of the island of Tongatapu in Tonga by driving on it with a GPS receiver turned on. He obtained an existing base map of the coastline, but no metadata were available to show the projection used. He chose the most likely coordinate system for the base map, but the road and the coastline were misaligned; on one side of the island, the road that the author had driven on appeared to be in the ocean.

A more general issue is that in many developing economies there is not much off-the-shelf geographic information available at the resolution needed to merge it with village- or community-level data. Information is more often available at
a coarser scale, making it difficult to link household locations to local geographic features. Even when information is available at high resolutions, it may not match up with the household survey because of differences in coordinate systems. It is therefore important for countries to have a spatial data infrastructure that coordinates collection activities so that different geographic data sets can be matched.13

Road Network Problems

Practical difficulties increase when constructing a road network for measuring distances or travel times. The algorithm used in a GIS for calculating the shortest distance requires good alignment between the lines and junctions of the digitized road network. For example, if digitized road segments at a junction do not line up, the algorithm will back-track and seek another path. These problems may be especially apparent when roads have been digitized for another reason, such as cartographic display, so once again metadata about the origin and purpose of geographic layers used in conjunction with GPS data are very useful. It is also sensible to budget considerable research assistance time to clean a road data set since misalignment problems are common. For example, the road network data set underlying the service areas for money transfer facilities in figure 1 took more than a week to clean. The digitized roads had been obtained from an earlier cartographic project, so even though it looked like a digital road network, it was more like a picture of a network and considerable work was required to convert it into the continuous lines and junctions needed for calculating travel distance and time.

One way to reduce the effort required to obtain a usable roads network data set is to digitize only the main roads and then to assume a network of feeder roads, which will automatically have nicely aligned junctions. This approach is used by Staal et al. (2002, 2003) who study market access and its effect on market participation and technology adoption for smallholder Kenyan dairy farms. To build a road network linking their sample of farms to Nairobi and other urban areas, they use topographic maps to digitize three classes of roads: all weather, bound surface; all weather, loose surface; and dry weather only roads. Since this classification leaves many of the surveyed farms off the actual road network, they add a 4 km grid of assumed feeder roads to fill in the areas between existing roads. It is not clear how much error is introduced by using this combination of actual and assumed roads.

Confidentiality Issues

The accuracy of GPS in measuring household and community locations also poses a challenge for maintaining the privacy of survey respondents. VanWey et al.
(2005) discuss the potential conflicts between the ethical need to ensure the confidentiality of information collected about human research subjects and the desire to link the characteristics and actions of individuals or households with a particular location. Uncertainty about how to proceed may mean that spatially explicit data are underutilized, undermining the role of data sharing and data preservation in advancing science or that researchers inadvertently disclose information that can identify survey respondents. These conflicts affect not only the original producers of data but also any data archivists charged with maintaining the database and providing it to other researchers while honoring the commitments to confidentiality made when the data were collected.

Moreover, this conflict between the confidentiality and the usability of GPS data is not limited to the sharing of data. It also affects the reporting and displaying of results based on geo-referenced data. For example, to show the confidentiality challenges posed by mapping point data, Curtis, Mills, and Leitner (2006) reverse address match back to individual residences (re-engineer), a newspaper map showing the locations in New Orleans where deaths occurred during Hurricane Katrina. Each location mark in the newspaper map covered approximately one and a half city blocks, and the map showed no roads and had few reference points. Nevertheless, over 30 percent of the re-engineered locations fell within 25 m of the actual residence where a death occurred. (Validation for the actual locations where the deaths occurred came from the search and rescue markings on dwellings, which were recorded during a field survey.) In several cases, the re-engineered location matched the field-verified residence where the death occurred. The authors scatter a series of random coordinates throughout the study area to show that chance alone would not give the same level of discovery. They conclude that “[t]he fact that many of the re-engineered coordinates could be used to identify an actual address, or an address within the immediate vicinity, should sound a note of caution for academics publishing maps displaying human cases as points” (p. 53).

Typical approaches for maintaining confidentiality are to allow access to GPS data only to approved researchers who promise not to identify respondents, to convert point data to surfaces or distances to avoid revealing individual locations, to aggregate and report data only for larger areas, and to use a geographic masking procedure that adds stochastic or deterministic noise to the geographic coordinates for sampled households and communities. Many surveys use a combination of these four approaches. Human-subject panels (which review the benefits and risks to subjects of research projects) can play an important role in protecting confidentiality, but researchers need to be aware of the costs of the different approaches.

Approval procedures are generally stricter for obtaining geo-referenced data than for ordinary household survey data. For example, researchers have to provide
additional justifications and commitments before they can obtain (masked) GPS
data on community locations in Demographic and Health Surveys. Stricter yet are
the rules for access to data enclaves. Data enclaves are usually located within a
survey organization, and accredited researchers are required to come to the
enclave to run their analysis. All output is checked for disclosure risk before
release, and there are typical restrictions on the linking of data sets and on the
identity and location of individual respondents. There are often entry fees to use
data enclaves, and the limited number of enclave locations may act as a barrier to
research. Remote (or virtual) data enclaves are being explored in some countries
to overcome these barriers. Rules on minimum cell sizes or the size of spatial units
to be mapped can ensure that no information identifying individuals is provided to
researchers. Cromley, Cromley, and Ye (2004) describe a system in which user
queries yield results only if the cell contains at least six records and the minimum
population in the smallest mapping unit is about 1,000.

Point data on household locations can be converted to a continuous surface to
represent the spatial distribution of characteristics or outcomes without identify-
ing respondents. For example, geographers have a variety of spatial interpolation
methods, such as spatial variants of the kernel density estimators increasingly
used by economists (Bithell 1990). Alternatively, point location coordinates can
be replaced by distances to various features of interest in any public-release data
set. However, these methods are not very flexible and are likely to limit future
research use of the data. Surfaces do not provide the micro-data needed for
studies that seek to measure causal impacts. The features of interest for which dis-
tances are reported are likely to vary from one study to another, and as distances
to more features are included, the possibility of using triangulation to identify
household locations increases. Moreover, distances are often needed for more than
features of interest. For example, knowing the position of households relative to
each other is important in studies of learning from neighbors (Conley and Udry
2005) and improves the modeling of spatial autocorrelation (Gibson and Olivia
2007).

Aggregating groups of observations into larger reporting units is widely used
for maintaining the confidentiality of survey and census records. With GPS coordi-
nates, the locations of individual households could be aggregated into larger areal
units such as census blocks or census tracts so that what is reported is an admin-
istrative code for the larger area or a polygon showing the boundaries of the area
where the household is located. These techniques can also be applied to the visual
display of data by using dot points on a map that are sufficiently large to prevent
disclosure risk. For example, VanWey et al. (2005) show how the size of the
required buffer around the locations of sampled schools in a U.S. survey would
need to vary from 6 km in a city to more than 50 km in the countryside to hold
disclosure risk to 5 percent when mapping the sample points. But aggregation
seriously degrades the analyses that can be conducted, since so much detailed spatial information is lost. For example, Fefferman, O’Neil, and Naumova (2005) provide an example of how areal aggregation can yield little benefit of additional privacy and large costs in impaired ability of statistical tools to analyze patterns of disease prevalence.

Geographic masking methods work by modifying the geographic coordinates linked to each household or community. Either random perturbations or affine transformations can be used. If the relationship of sample points to each other is important, whereas the relationship to another data layer (say, a base map or a road network) is not, then simply moving all points by a given distance and direction or rotating them about a chosen point may preserve confidentiality and not greatly degrade the usability of the data. In a variant of this approach, the Rural Investment Climate Survey in Indonesia rotates points in each sample cluster. Normally, however, point locations obtained with a GPS are merged onto other data layers so these affine transformations will introduce error that reduces the usability of the GPS data.

Another option is to introduce random perturbation around the original point with the radius of the perturbation circle chosen by the data custodian, possibly weighting the size of the circle by population density at each point to take into account the effect of population density on the risk of disclosure (Kwan, Casas, and Schmitz 2004). For example, many recent Demographic and Health Surveys include information on HIV infection. Because of confidentiality issues, up to 2 km of random error in any direction is added to cluster locations in urban areas and up to 5 km in rural areas. Additionally, one point in each survey is displaced up to 12 km in any direction. Only limited research has been conducted on the effect of random perturbation on disclosure risk or the accuracy of results. Kwan et al. (2004) show that the accuracy of results diminishes as the size of the perturbation circle increases. Zimmerman and Pavlik (2006) show that releasing metadata about perturbation methods and having different masked versions of the same data set (for example, a spatially aggregated data set and a randomly perturbed data set) can considerably increase disclosure risk.

Conclusions

The removal of selective availability and the falling costs of GPS receivers have made the collection of GPS information increasingly feasible in household surveys, yet many household surveys still do not use this technology. This article argues that the collection of GPS coordinates should become a routine part of household surveys, since doing so can lead to better economics and better policy advice. In particular, the article has shown how GPS is being used to better
measure and understand the causal impacts of policies, policy externalities, and access to services and how using GPS can improve the quality of the data collected.

One of the greatest arguments to collect GPS information is the option value it gives for unforeseen future applications.\textsuperscript{15} As the stock of geo-referenced data increases within developing economies, a number of innovative applications are likely to emerge, which combine household survey data with other geographic information. In addition, interesting new research questions in econometrics and sampling methodology will emerge from the use of GPS. For example, typical household surveys often involve population-based clusters, which are not randomly spread across geographic areas (Montana and Spencer 2004). As a consequence, more research is needed to determine how best to estimate spatial models using a sample that is not spatially representative at the local level and how best to sample within communities to allow both spatial and nonspatial uses of the data.

The increasing prevalence of GPS data will have greater research value if current practices are improved. For example, the provision of accurate, clear and timely metadata about the data layers in existing GIS collections would allow more seamless and reliable merging of such data with GPS data. Ensuring respondent privacy while maximizing the research value of GPS data may be best achieved by having surveys approved by human-subject panels and by releasing data only to valid researchers who sign confidentiality provisions.

\section*{Funding}

Financial support from Marsden Fund grant UOW0503 is gratefully acknowledged.

\section*{Notes}

John Gibson is a Professor of Economics at the University of Waikato; his email address is jkgibson@mgmt.waikato.ac.nz. David McKenzie (corresponding author) is a senior economist in the Development Research Group at the World Bank; his email address is dmckenzie@worldbank.org. The authors are grateful to Kathleen Beegle, Chris Bennett, Piet Buys, Geua Boe-Gibson, Alan de Brauw, Uwe Deichmann, John Hoddinott, Ben Olken, Duncan Thomas, John Thyne, three anonymous referees, and participants at the 2007 Stanford Institute for Theoretical Economics Conference for helpful comments and advice. This article is dedicated to the memory of Piet Buys, who championed global information system use at the World Bank through his knowledge, enthusiasm, and eagerness to help others use this emerging technology.

1. \url{www.garmin.com/support/faqs/faq.jsp?faq=582&webPage=Main%20web%20page}

2. Specifically, a "base station" GPS receiver is set up on a precisely known location and used to compare position based on the satellite signals with this known location. The difference is then
applied to other GPS receivers in the area to correct their calculations of their unknown locations.

3. The Garmin eTrex GPS unit, used in a number of household surveys, could be purchased for $88 at Walmart.com and $93 at Amazon.com on February 21, 2007.


5. www.xyzmaps.com/NewPostcode.htm

6. Recent reviews of the use of GIS in economics that are based largely on developed countries are Overman (forthcoming) and Bateman et al. (2002).


8. Of course, researchers can use geography to create instruments without using GPS, through manual map work. Recent examples include Woodruff and Zenteno (2007), who used distance from the capital of the state in which an individual was born to the nearest station on the north–south railway lines as they existed in the early 1900s as an instrument for migration in Mexico. Hoxby (2000) used the number of stream mouths in a metropolitan area in the United States as an instrument for the number of school districts in examining the impact of school choice. GPS can make such applications more accurate and less time-consuming.

9. However, note that this application involves using distance within Tonga as an instrument for migration to another location (New Zealand). As a result, one is less concerned in this application about other geographic features in Tonga affecting the outcome of interest, since this outcome is in New Zealand.

10. See Anselin (2002) for an accessible review of spatial econometrics. Note also that in the agricultural example given here, the omitted climate and soil variables are likely to be correlated with the regressors of interest, so one will wish also to include detailed spatial variables as controls in the regression.

11. Roberts et al. (2006) report on a survey in Bukoba, Tanzania, where self-reported distance was compared with distances calculated by using pedometers and an estimate of average step length. They find that more than 60 percent of self-reported distances were more than twice the calculated distances.

12. A degree of latitude is 110.6 km at the equator and 111.7 km at the poles. A degree of longitude is 111.3 km at the equator, 55.8 km at 60 degrees latitude and only 16.9 km at 80 degrees latitude. Instead, the great circle distance should be used (http://mathforum.org/library/drmath/view/51711.html). Stata code that implements this formula is available in the “globdist” ado files written by Ken Simons (available through Stata’s online search function).


14. For more details see www.measuredhs.com/topics/gis/methodology.cfm


References


Zimmerman, Dale, and Claire Pavlik. 2006. “Quantifying the Effects of Mask Metadata Disclosure and Multiple Releases on the Confidentiality of Geographically Masked Health Data.” Iowa City: University of Iowa Department of Biostatistics.
STAY ALERT TO THE VERY LATEST RESEARCH

DESIGN YOUR OWN EMAIL ALERTS TO TRACK JOURNAL CONTENT THAT IS IMPORTANT TO YOU, USING CITETRACK

Track topics
Tell CiteTrack which words or subjects to watch for in new content

Track authors
Be alerted whenever key authors you are following publish a new paper

Track articles
Know whenever a paper of interest to you is referenced by another paper or letter

CiteTrack checks for content matching your criteria in new issues of the journal you’ve asked it to monitor, and emails you each time new content is found.

With around 1000 journals currently available through this service, you can easily monitor a wealth of quality research with one simple service.

Sign up for CiteTrack today – go to your favourite Oxford Journals title, and click on the CiteTrack link under ‘Alerting Services’ to set up your custom alerts.
STAY ALERT TO THE LATEST CONTENT FROM OXFORD JOURNALS... FOR FREE!

WE WILL AUTOMATICALLY EMAIL YOU THE FORTHCOMING TABLES OF CONTENTS FOR NEW ISSUES OF JOURNALS PUBLISHED BY OXFORD JOURNALS

Go to this journal’s homepage and click on the link for ‘email table of contents’ under ‘Alerting Services’.

www.oxfordjournals.org
ACCELERATED ONLINE PUBLICATION

READ PAPERS FROM THIS JOURNAL ONLINE, WEEKS IN ADVANCE OF SEEING THEM IN PRINT

The advantage for readers:
Read the very latest research online ahead of papers appearing in a journal issue.

The advantage for authors:
See your paper receiving citations faster than ever before.
Visit www.oxfordjournals.org, select this journal from the list to reach its homepage, and click on Advance Access. Bookmark the page and regularly check for the latest accepted papers.

www.oxfordjournals.org