Second generation fiscal federalism: The implications of fiscal incentives

Barry R. Weingast 1

Stanford University, Department of Political Science, Encina Hall, Stanford, CA, United States

A R T I C L E   I N F O

Article history:
Received 18 January 2007
Revised 7 November 2008
Available online 25 December 2008

JEL classification:
H77
H11
H71
R11
R58

Keywords:
Fiscal federalism
State and local taxation
Intergovernmental relations
Revenue

A B S T R A C T

First generation fiscal federalism (FGFF) studies the performance of decentralized systems under the assumption of benevolent social planners. Second generation fiscal federalism (SGFF) studies performance based on the fiscal and political incentives facing subnational officials. The paper focuses on three aspects of SGFF. First, it considers the design of intergovernmental transfers. While FGFF emphasizes correcting vertical and horizontal equity, SGFF emphasizes the importance of fiscal incentives for producing local economic prosperity. SGFF extends FGFF approaches by showing how non-linear transfer systems can produce both equalization and high marginal fiscal incentives to produce local economic growth. Second, the paper raises the fiscal incentive approach, showing how different tax systems produce different fiscal incentives for political officials to choose policies. Third, the paper discusses the interaction of democracy and fiscal federalism.

© 2008 Elsevier Inc. All rights reserved.

[Second generation fiscal federalism represents] a new literature on fiscal federalism that examines the workings of different political and fiscal institutions in a setting of imperfect information and control with a basic focus on the incentives that these institutions embody and the result behavior they induce from utility-maximizing participants (Oates, 2005, p. 356).

Much fiscal analysis of developing countries is on the following pattern: the academic literature is drawn on to construct a model fiscal system; the existing situation in a particular country is examined to determine how it diverges from the model; and a fiscal reform is then proposed to transform what is into what ought to be…. In contrast, my approach is first to study in detail exactly how the existing system works, and why it works that way, in order to have a firm basis for understanding what changes may be both desirable and feasible. My emphasis has thus always been more on what can be done than on what should be done (Bird, 1992, x, emphasis in original).

1. Introduction

Why do federal nations exhibit such widely different economic performance? Some are rich (Switzerland and the United States) while some are poorer (Argentina and Brazil); some exhibit fast-paced growth (modern China) while others little growth (Mexico). In this essay, I explore this issue by surveying the new literature on second generation fiscal federalism (SGFF), which complements first generation fiscal federalism (FGFF), The distinction between FGFF and SGFF parallels that made by Musgrave (1959, p. 4):

[Theories of Public Economy] can be approached in two ways. First, we attempt to state the rules and principles that make for an efficient conduct of the public economy…. In the second approach, we attempt to develop a theory that permits us to explain why existing policies are pursued and to predict which policies will be pursued in the future.

FGFF is largely normative and assumes that public decision-makers are benevolent maximizers of the social welfare (Musgrave, 1959; Oates, 1972; Rubinfeld, 1987). SGFF builds on FGFF but assumes that public officials have goals induced by political institutions that often diverge from maximizing citizen welfare (Oates, 2005; Garzarelli, 2004; Qian and Weingast, 1997; see also Brennan and Buchanan, 1980; Salmon, 1986; and Wicksell, 1967). As Hatfield (2006) puts it, “Economic policy is not decided by benevolent social planners, but by government officials, usually with at least one eye to their reelection prospects.”

The distinction should not be overdrawn – no clean demarcation exists between the generations, and many first generation works develop considerable positive implications. Nonetheless, the
distinction is important because it emphasizes the extension of normative fiscal federalism to take systematic account of public official incentives. SGFF models provide a range of new insights into fiscal federalism (Oates, 2005). This approach also provides new normative prescriptions for the design of federal systems, including how many of the prescriptions of FGFF should be adapted given more realistic political choice environments. SGFF explores how various institutions align – or fail to align – the incentives of political officials with those of citizens. This approach is central to understanding differential federal performance.

In this essay, I survey a range of SGFF ideas and explore their implications for developing countries in the context of decentralization and democratic governance. I begin with the perspective of market-preserving federalism. By studying the conditions and incentives of subnational government authority and policymaking, this perspective provides a comparative theory of federal performance: federal systems that satisfy different combinations of the market-preserving federalism conditions differ in predictable ways. The comparative analysis helps explain why decentralized systems exhibit so much variance in behavior. This analysis allows us to study a range of “pathologies of federalism” – fiscal institutions that produce perverse or market-distorting outcomes.

I next discuss SGFF implications for intergovernmental transfer systems. FGFF models emphasize the importance of transfers for mitigating vertical and horizontal imbalances. The SGFF approach emphasizes the importance of incentives generated by local tax generation for fostering local economic prosperity. Subnational and urban governments are more likely to provide market-enhancing public goods when they capture a large portion of the increased tax revenue generated by greater economic activity. SGFF approaches have significant implications for the design of transfer systems so that equalization goals can be achieved while providing public officials with incentives to foster thriving local economies.

Next, I turn to the fiscal incentives approach with long roots in the study of fiscal federalism.2 The idea is that, whatever their goals, public officials favor policies that relax their budget constraints. Different systems of taxation and intergovernmental transfers therefore directly affect local governmental behavior and policy choice. I study several types of fiscal incentives associated with intergovernmental transfer systems, including the design of markets and corruption.

I also discuss the role of democracy. Democracy is a source of freedom and expression for citizens, and when it works well, it provides citizens a means to express choices and to hold public officials accountable. But democracy often fails in practice for developing countries. I show how the fiscal system affects the performance of democracy by investigating one source of failure of democracy called “tragic brilliance,” the idea that voting can create social welfare.3

3 Inman and Rubinfeld, 1997b).3 Others working in the context of developing countries, follow Bird’s (1992) point noted in the paper’s headnote. A large body of work studies various forms of common pool problems, of which three stand out: the so-called “race to the bottom”; problems with a soft budget constraints for subnational governments; and common pool problems associated with centralized provision of local public goods. Still other scholars call for understanding differences among federal systems in order to learn what institutions support market-preserving (Weingast, 1995). Beginning with Riker (1964), another strand in the literature emphasizes the political aspects of federal performance, particularly political parties. Relatedly, scholars investigate the self-enforcing rules necessary maintain federal stability.

2. Market-preserving federalism and the comparative theory of decentralized governance

Local governments exist within a complex set of institutional arrangements, with political, legal-constitutional, and economic aspects. This section develops a framework for analyzing how different institutional arrangements affect the performance of local governments.

Federalism, and decentralization more generally, encompasses a wide range of different political-economic systems, not one, whose political and economic properties vary widely (Shah, 1997b; Watts, 1999). As Litvak et al. (1998, p. vii) observe, “decentralization is neither good nor bad for efficiency, equity, or macroeconomic stability; but rather that its effects depend on institution-specific design.” We therefore cannot speak of the tendencies of federalism per se. Some federal systems promote macroeconomic stability and economic growth while others just the opposite.

Consider: For the last three centuries, the richest nation in the world has almost always been federal. The Dutch Republic from the late sixteenth through mid-seventeenth centuries; England from the late seventeenth or early eighteenth and mid-nineteenth centuries (a de facto though not de jure federal system); and the United States from the late nineteenth to the present. Similarly, modern China, a de facto federal state, has experienced sustained rapid growth. India, having grown slowly for several decades, has experienced high growth in the last. In contrast, the large Latin America federal states of Argentina, Brazil, and Mexico, have all fared much more poorly. How do we account for such large differences in economic performance?

In this section, I summarize a comparative theory of decentralized governance that explains the differential economic performance of various types of decentralization. This framework helps understand some of the institutions necessary to support decentralization that provides political officials with incentives to improve social welfare.
The comparative theory of federal performance begins with a set of conditions that differentiate federal systems.11 All federal systems decentralize political authority, so a necessary condition for federalism is:

(F1) Hierarchy. A hierarchy of governments exists with each level having a delineated scope of authority.

Yet federal systems differ enormously in terms of the policy authority assigned to different levels of government. The following conditions characterize how federal states assign authority among national and subnational governments.

(F2) Subnational autonomy. Subnational governments have primary both local regulation of the economy and authority over public goods and service provision.

(F3) Common market. The national government provides for and polices a common market that allows factor and product mobility.

(F4) Hard budget constraints. All governments, especially subnational ones, face hard budget constraints.

(F5) Institutionalized authority. The allocation of political authority is institutionalized.

We can characterize different federal systems by which of conditions they satisfy, ranging from the hierarchy condition alone to all five conditions.12

An ideal type of federalism, called market-preserving federalism, satisfies all five conditions (Weingast, 1995). These conditions make explicit some of the political assumptions implicit in FGFF. Indeed, many of the major results in this approach implicitly assume most or all these conditions, including Oates’s (1972) “decentralization theorem,” Tiebout’s (1956) interjurisdictional competition, and Musgrave’s (1959) solution to the assignment problem.

Scholars of fiscal federalism have long argued that federalism places subnational governments in competition with one another (Tiebout, 1956; Oates, 1972; Brennan and Buchanan, 1980). Competition gives subnational governments the incentive to foster local economic prosperity rather than costly market intervention, service to interest groups, and corruption. Competition among jurisdictions limits a subnational government’s ability to abuse its policy authority, for example, by precluding on investments or by granting privileged positions, such as monopolies or above market wages to government workers. Governments that fail to foster markets risk falling land values and the loss of capital and labor – and hence valuable tax revenue. Put another way, interjurisdictional competition provides political officials with strong fiscal incentives to pursue policies that provide for a healthy local economy.

Effective inter-jurisdictional competition requires several institutional conditions. First, subnational governments must have the authority to adapt policies to their circumstances; hence, the subnational autonomy condition (F2). Consistent with the FGFF assignment principle, these governments must have considerable power to regulate local markets, to tailor the provision of local public goods and services to local circumstances, and to set tax rates, ideally to reflect local demand for public services (Musgrave, 1959; Oates, 1972).

Many federal systems restrict the policy authority and independence of subnational governments, compromising the benefits of federalism. Examples include Mexico throughout much of the late twentieth century; India from independence through the mid-1990s; and Russia under Putin.

Second, effective competition among jurisdictions requires product and factor mobility across jurisdictional boundaries—hence the common market condition (F3). This condition has held for the United States since the inception of the Constitution. Indeed, rising trade barriers among the states was a Federalist argument against the Articles of Confederation. In contrast, India allows internal trade barriers, and Russia restricts the movement of labor, capital and goods across regional borders in various ways.

The failure of the common market condition creates a pathology in which subnational government becomes a de facto “national government” within its jurisdiction. By reducing the penalties for costly market intervention, rent seeking, and corruption, internal trade barriers short-circuit interjurisdictional competition and hence federalism’s constraints on subnational policymaking. Because many developing federal systems limit factor mobility, particularly labor mobility, Bardhan (2002) questions whether the standard FGFF framework is relevant for many developing countries.

Third, effective interjurisdictional competition requires a hard budget constraint (F4), which concerns both government borrowing and fiscal transfers among levels of government.13 This condition requires that subnational governments bear the full financial consequences of their policy decisions, so that they cannot spend beyond their means or endlessly bail out failing enterprises. A hard budget constraint also precludes the national government from bailing out subnational governments that go into deficit, whether through cash transfers or forgivable loans.

FGFF logic shows that a hard budget constraint provides local political officials with incentives for prudent fiscal management of their jurisdiction. As Shah (1997c) concludes, “to ensure fiscal discipline, governments at all levels must be made to face financial consequences of their decisions.” In contrast, subnational governments facing a soft budget constraint have incentives to spend beyond their means, pursue costly market intervention, provide costly benefits to interest groups, endlessly subsidize ailing enterprises, and engage in corruption. The expectation of bailouts lowers the financial costs to the subnational governments (though not to the country) of these expenditures. Argentina in the 1980s and Brazil in the 1990s both experienced hyper-inflation as their provincial governments spent without limits, forcing the federal government to bail them out.

The final condition—institutionalized authority (F5)—provides the glue for the decentralized system. This condition requires that decentralization must not be under the discretionary or unilateral control of the national government (similarly, local decentralization within a region must not be solely at the discretion of the regional government). Instead, a set of institutions must exist that prevent the national government from altering or undoing aspects of subnational autonomy. In the absence of this condition, the national government can compromise subnational government autonomy and hence the benefits from competition among them. The Mexican president, for example, has historically had the power to remove governors (Carlos Salinas, President of Mexico from 1988 to 1994, removed over half the governors during his six year term). This power dramatically reduces the independence of the states because the federal government can threaten those states which do not conform to the

---

11 This section draws on McKinnon (1997), Montinola et al. (1995) and Weingast (1995, 2005); Inman and Rubinfeld (1997a) provide alternative sets of conditions for differentiating among federal systems.

12 To make this discussion manageable, I ignore many subtleties and simply assume that each condition either holds or not. For further details, see Montinola et al. (1995) and Weingast (1995).

13 Several works provide excellent discussions of the HBC, especially the specific institutional necessary to implement it. See Dillinger and Webb (1999), Haggard and Webb (2004), McKinnon (1997), Rodden (2005), Rodden et al. (2001), and Wildasin (1997).
The power of the central government to intervene in state affairs differs dramatically across federal systems. The Indian Constitution grants the central government relatively unconstrained power to take over states (although interventions by the courts have changed this in recent years), whereas the Spanish Constitution places a complex set of constraints on the central government's ability to take over a state (Stepan, 2004a).

The institutionalized authority condition is easy to understand in the abstract, yet we know too little about the mechanisms that make some federal systems succeed. A host of writers follow Riker (1964) and argue that the form of the party system is essential to maintaining federalism. Some party systems allow national elites to dominate the parties; others allow local elites to dominate; and still others allow for a balance of power among national and local elites. When national elites dominate parties, they are likely to force local leaders to accept institutional changes that compromise local government powers (as in Mexico under the PRI, 1940–2000, India under the Congress Party, 1950–1989, and Russia under Putin, 2000–present). In contrast, a party system dominated by local elites is more likely to force national elites to accept subnational government common pool behavior, such as bailing out subnational deficits (as in Brazil in the late 1990s). Finally, a party system balanced between national and local elites is more likely to support decentralization, as both local and national elites guard their own prerogatives (as in the United States). This perspective begs the issue of what creates different types of party systems.

Market-preserving federalism limits the exercise of corruption, predation, and rent-seeking by all levels of government. This form of decentralization is potentially important for developing countries, where central government market-intervention frequently bestows many sectors with monopolies and various forms of protection from competition. Market-preserving federalism limits the ability all levels of government to create monopolies and mass state-owned enterprises whose primary political purpose is to provide jobs, patronage, and other forms of inefficient economic intervention that plague developing countries (as Brennan and Buchanan, 1980 have observed). A subnational government that seeks to create monopolies, engage in extensive corruption, or arrange a privileged position for an interest group places firms in its jurisdiction at a disadvantage relative to competing firms from less restrictive jurisdictions.

The comparative federalism framework also characterizes a set of pathologies of federalism, forms of market-distorting federalism that fail to provide incentives to foster and preserve markets (see Wibbens, 2005, ch. 2). To summarize, the absence of one or more of the conditions (F2)–(F5) implies some form of inefficiency or pathology.

- The absence of subnational policy authority (F2) inhibits the subnational competitive process and the ability of subnational governments to tailor policies to local conditions.
- The absence of a common market (F3) directly hinders competition among jurisdictions, so that subnational governments are more likely to engage in corruption, rent-seeking, and inefficient resource allocation. Restrictions on factor mobility have a similar effect.
- The absence of a hard budget constraint (F4) allows subnational governments to live beyond their means so that they engage in more corruption, non-remunerative benefits to interest groups, and endless subsidies to inefficient enterprises.
- Finally, the absence of institutionalized authority (F5) allows the center to threaten subnational jurisdictions who seek policy independence.

This brief analysis of federal pathologies suggests why many recent decentralization reforms fail. Because decentralization so often does not satisfy one or several of the market-enhancing conditions, it fails to provide subnational governments with incentives to foster markets. Indeed, too frequently in the developing context, decentralization involves very limited local government policy or tax independence. For example, Thomson (2006) observes that much devolution of power in Africa often grants too little policy authority, involves too many unfunded mandates, and grants subnational governments only unproductive taxes and inadequate and unpredictable intergovernmental transfers. Similarly, Wiesner (2003, pp. 17–19) criticizes Bolivia's decentralization because it granted subnational governments “insufficient capacity (incentives)” or policy authority while allowing a soft budget constraint based on debt relief.

A related problem in the developing world is that decentralization in a truly predatory state is not likely to succeed. A central government that is not committed to decentralization has numerous ways to undermine subnational government performance, including inadequate revenue, constraints on subnational policymaking and unfunded mandates, and direct threats to political officials who deviate from the central government's policies.

Let me conclude the discussion of comparative federalism with a contrast between post-reform China and India under the Congress party (1950–1990) that helps characterize the differential performance of these two federal systems. China is a market-preserving federal system embedded within an authoritarian regimes (Jin et al., 2005; Montinola et al., 1995). Provinces (and lower governments, including townships and villages) have been the great engine of economic growth. Since 1980, subnational governments have exercised wide ranging policy and fiscal independence, allowing them to innovate pro-market policies and provide market-enhancing public goods. China satisfies nearly all the conditions, with the possible exception of F5: Provinces have substantial policymaking authority (F2) and face a relatively hard budget constraint (F4), although there remain some trade barriers (F3) and the institutional security of the system (F5) remains in some doubt (see discussion below).

In contrast, India's sluggish performance from independence through the early-1990s reflected its centralized federal system, where the central government made most of the important policy decisions (compromising F2) and imposed some restrictions on the movement of goods across states (F3). States did face a hard budget constraint (F4). However, as discussed below, the center also had the ability to take over states, and it used this discretion in part to undermine successful opposition parties. This political predation compromised state independence (both F2 and F5). Centralized federalism, with too little state policy or fiscal independence, prevented states from innovating and fostering more market-enhancing local economies than that favored by the center. Importantly, as the center has loosened the constraints on states, states have become more innovative, and India's economic growth has improved significantly.
3. The fiscal incentives approach: taxation and the design of transfer systems

The fiscal incentives approach emphasizes how fiscal institutions create incentives for subnational political officials that affect their policy choice and hence their jurisdiction's performance. Whatever the goals of subnational officials, greater revenue relieves their budget constraint, allowing them to further their goals. Political officials of all stripes are therefore biased toward policies that increase their revenue, allowing them to finance more activities. This assumption differs from revenue maximization of Leviathan (Brennan and Buchanan, 1980). Leviathan is an extreme form of fiscal interest in which officials have no goals other than revenue maximization. The fiscal incentives approach is more general, allowing political officials to goals other than revenue; but they all care about revenue because it allows them to pursue their other goals. This observation implies that the fiscal system directly influences whether governments choose market-fostering or distorting policies.

Economists have long known about this principle, although they have not always studied it systematically. Tiebout (1956), for example, discussed the beneficial effects of the property tax for local government. Because the value of public goods is capitalized into the value of local property, dependence on property taxation leads city managers to choose public goods that maximizing local property values. Moreover, city managers facing intense interjurisdictional competition have incentives to maximize property values as a means of inducing scarce capital and labor to locate and remain in their jurisdiction. Because these taxes provide general incentives for local political officials to design policies that foster markets and attract capital and labor, property taxes are an important component of local government fiscal structure.

3.1. Transfer systems, vertical and horizontal equalization, and incentives

The fiscal incentives approach has significant implications for the design of transfer systems within federal systems. The FGFF rationale for intergovernmental emphasizes vertical and horizontal tax imbalances, spillovers of benefits, and forestalling costly tax competition (see Bahl and Flatters, 1982; Boex and Martinez-Vazquez, 2006; Oates, 1972). Vertical imbalances arise when the center collects taxes more easily and at lower economic cost than subnational governments; it also arises when the central government preempt subnational government revenue sources (McLure, 1993). Efficiency considerations suggest that the center raise more taxes and then transfers funds to subnational government to finance a portion of their expenditures. Horizontal imbalances arise because regional economies differ in their income and hence in their ability to provide citizens with public goods and services. Here too transfers from the center can mitigate these imbalances by providing greater funds to poorer localities.

SGFF models emphasize the importance of revenue generation by subnational governments (Rodden, 2003; Singh and Srinivasan, 2006; Careaga and Weingast, 2003); subnational governments that raise a substantial portion of their own revenue tend to be more accountable to citizens, to provide market-enhancing public goods, and to be less corrupt.

Many FGFF scholars recognize this principle. Shah (1997a, 1997b) argues this point in a series of influential papers. McLure (1998, p. 1) observes that “Subnational governments that lack independent sources of revenue can never truly enjoy fiscal autonomy; they may be — and probably are — under the thumb of the central government.” Similarly, Bahl and Linn (1992, p. 428), in their authoritative study of local fiscal federalism in developing countries, observe that “grants can make local governments less accountable for their fiscal decisions (they may now increase spending without increasing taxes); hence there will be less incentive to improve the efficiency of local government operations and develop innovative methods of delivering public services.”

Despite these observations, FGFF analyses of intergovernmental transfers tend to focus on equity considerations rather than emphasizing the incentive effect of transfers on subnational government policymaking or growth. As Singh and Srinivasan (2006, p. 34) observe:

The standard public finance question takes subnational jurisdiction’s income as given and looks at the incentive effects of tax assignments and transfers. The [SGFF] growth perspective examines the effects of the tax and transfer system on incentives to increase income (e.g., through public or private investment).

Singh and Srinivasan further suggest that “the allocative efficiency of the tax system in a standard public economics sense is of second order importance relative to fiscal autonomy on the revenue side” (23).

The SGFF logic provides two related reasons for these conclusions. First, transfers that are negatively related or only weakly positively related to subnational income growth give local governments poor fiscal incentives to foster local economic growth. Second, such transfer systems induce greater corruption and rent-seeking. This and the next subsection studies the first issue, while the following studies the second.

The attempt to correct vertical and horizontal imbalances in developing countries often means that these transfer systems exhibit poor responsiveness to localities that foster local economic growth. For example, the Finance Commission’s transfers of revenue to states in India reflect a series of weights for different criteria: 62.5 percent is negatively related to a state’s income, so that poorer states receive greater funds; 10 percent on the basis of population; and the remainder somewhat evenly divided among state area, an index of infrastructure, tax effort, and fiscal discipline. This type of intergovernmental transfer system provides poor fiscal incentives for subnational jurisdictions to foster local economic growth: most of an increase in local revenue goes to the center (Singh and Srinivasan, 2006).

To analyze these incentives, consider a transfer system, such as the Indian Finance Commission’s, set by formula that takes into account various economic and demographic characteristics, such as income and population. Suppose that the formula is fixed with reference to a given year so that the center allocates revenue using the same proportions each year, with the only variable across years being the size of the revenue pool to be divided among subnational governments.

If there are n provinces, then the average province receives 1/n of the total revenue pool, no matter how good or bad its policies. Let the total revenue pool be R, so that the average province receives a share $S = R/n$ of the total pool. Now let the province alter policies to foster local economic growth so that the revenue generated from the province, r, increases. The average province’s share increases by $\partial S/\partial r = 1/n$. In other words, the province receives

---

36 Bahl and Linn (1992, especially chs. 4–6) provide one of the most comprehensive discussions of the property tax in decentralized systems. Fischel (2001) and Glaeser (1996) explain why property taxation leads local governments to focus on citizen welfare. See also Hotby (1999). Nonetheless, these incentives are incomplete, as Apple and Romer (1991) have shown.

17 These figures are for the 11th Finance Commission. See Rao and Singh (2005, ch. 9, especially table 9.3) and Singh and Srinivasan (2006). The Planning Commission also transfers money to states based on different criteria.
of the total increase in revenue generated solely from its increased investment in the local economy. The province bears the full expenses for the market-enhancing public goods but captures only 1/11 of the fiscal return.

Careaga and Weingast (2003) called the poor incentives of these transfer systems “fiscal law of 1/11.” In a country with even a modest number of states, this proportion is quite small. One of 23 provinces in Argentina would receive only four centavos for each newly generated peso in taxes; while one of 33 states in Mexico would receive three centavos. In contrast, fiscal systems that allow growing regions to capture a major portion of new revenue generated by economic growth provide far stronger incentives for local governments to foster local economic growth.

No systematic study exists of these fiscal incentives, but a few investigations calculate the proportion of local revenue captured by local governments in particular cases. Although it is unlikely that this single variable accounts for long-term economic growth, the results suggest an interesting pattern for developing countries (see Table 1). Careaga and Weingast (2003) calculated the marginal revenue retention rate for Mexico for different periods. In 1995, for example, the rate was 23.3 percent. But there were periods when all state revenue was put in a common pool and then divided by a sharing rule, which meant that the percentage for the average state was close to 1/33 (for 33 states). Zhuravskaya (2000) calculated that this figure as 10 percent for Russian cities.

Table 1

<table>
<thead>
<tr>
<th>Degree of revenue capture</th>
<th>Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very high</td>
<td>High</td>
</tr>
<tr>
<td>States in U.S., 19th century</td>
<td>High</td>
</tr>
<tr>
<td>Provinces in China, 1982–1993</td>
<td>High</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>China, immediate pre-reform era</td>
<td>Low</td>
</tr>
<tr>
<td>Mexican states, 1980–1995</td>
<td>Low</td>
</tr>
<tr>
<td>Indian states, 1950–1990</td>
<td>Low</td>
</tr>
<tr>
<td>Russian cities &amp; regions, 1990s</td>
<td>Low</td>
</tr>
</tbody>
</table>

Sources: China (Jin et al., 2005), Mexico (Careaga and Weingast, 2003), India (Rao and Singh, 2005), and Russia (Blanchard and Shleifer, 2000; Zhuravskaya, 2000).

Transfer systems may exhibit other perverse fiscal incentives. Some systems are explicitly “gap-filling,” meaning that provinces with larger deficits receive larger transfers. Because these systems subsidize spending beyond revenue, they provide subnational governments with incentives to spend beyond their means. Gap-filling has long been a feature of transfers within Indian system, a problem that has gotten considerably worse in the last decade. Relatedly, both Argentina and Brazil in the late twentieth century had local branches of the central bank that essentially allowed provinces to transfer debt to the central government, leading to massive financial problems.

Another problem is that some transfer systems fail by design. Wiesner (2003, p. 23) argues that decentralization in Latin America often emphasizes subnational government entitlements to revenue rather than markets and incentives: “These frameworks tend to neglect market-based mechanisms and make the capture of large unconditional transfers an easy ride for public sector rent-seeking.” For example, Bolivia, Brazil, and Ecuador considerably increased the transfer of revenue without increasing the policy responsibility, allowing subnational governments to use these funds for patronage rather than local public goods (Wiesner, 2003). Common pool problem of budgets have also plagued this region (Jones et al., 2000; Stein, 1998).

Courchene (1981) and McKinnon (1997) raise a related incentive problem with transfer schemes that are designed to provide substantial subsidies to the poorest regions in rich countries. McKinnon, for example, contrasts the huge subsidies by Canada of the Eastern Maritime Provinces and by Italy of Mezzogiorno in Southern Italy with the lack of subsidies by the United States to the American South. McKinnon suggests that the revenue transfers in Canada and Italy create dependency and a soft budget constraint. Transfers allow these regions to finance ailing and inefficient enterprises, seeming to saddle Southern Italy with highly capitalized, loss-making enterprises. The regional economy is far less likely to adapt so that it becomes more like the vibrant national economy. In contrast, southern states in America faced a hard budget constraint and no national subsidies. The poorest region in the United States after the Civil War through mid-20th century, southern states were able to grow rich by redesigning their economies with low regulatory burdens relative to the industrialized North and to take advantage of lower labor costs. This adaptation fostered the booming sun belt economy of the late twentieth century. McKinnon argues that the economic rise of the American South is unlikely to have occurred had it been subsidized in the manner of the Canadian Maritimes and the Italian Mezzogiorno. For this reason, Courchene argues that these types of regional transfers are self-perpetuating.

3.2. SGFF implications for the design of transfer systems

SGFF logic suggests that the design of transfer systems should take at least two costs into account. Following FGFF logic, these should lower the tax burden on the economy and limit tax competition; and following the SGFF fiscal incentive approach, transfer systems should reward subnational governments that foster local economic growth. The above discussion shows that many transfer systems achieve equalization at the expense of subnational government incentives to foster economic prosperity. Yet this tradeoff is neither a necessary nor inevitable. In particular, it was not true of China’s fiscal system from 1982–1993.

Federations can simultaneously achieve all three goals – horizontal equalization, preventing tax competition, and ensure high marginal fiscal incentives – by designing transfer systems with non-linear functions. The poorest provinces with only limited ca-

---

18 Following the results of Weingast et al. (1981) “law of 1/11”; see also Inman (1988).
19 The data presented in Shah (1998, pp. 136–144) suggest that the figures for Pakistan are less than a third. The data in Rao and Singh (2005, ch. 9) suggest that the figures in India are similar.
20 The opposite phenomenon is equally problematic – the devolution of considerably more authority and responsibility without the fiscal resources to implement it.
21 Krueger (2006) uses similar logic to explain the difference between the vibrant economy in Poland just east of border with Germany and the lackluster economic performance of the former East Germany just west of the border: massive transfers from the German government have deterred economic development.
capacity to grow or to tax should be treated in a manner similar to the existing transfer systems. For other provinces: (1) the center keeps track of its revenue collection by province; (2) a step function allows the center to capture a moderate or moderately high proportion of revenue generated from a province up to a revenue level fixed in advance and then allows the province to keep a high or very high proportion of all revenue raised above that level. The step function provides the province faces high marginal incentives to foster local economic prosperity.

A central government using a traditional scheme might keep 75 percent of all revenue, whereas one using a non-linear scheme might keep 80 percent up to some amount and then 25 percent thereafter (with the breakpoint chosen so that the central government initially captures the same amount of revenue under two schemes). Suppose a province creates conditions for growth of 10 percent per year for five years under the two schemes. Under the non-linear scheme, the province keeps a portion that is three times greater than under the traditional scheme so that it has far stronger fiscal incentives to foster growth.

The advantage of the high marginal transfer system is that taxpayers are more likely to bear the expenses of market-enhancing public goods when they receive a large fiscal return. Moreover, this transfer system is Pareto improving. Although some provinces will get richer than others, the total amount retained by the center is larger than if these provinces had grown less so that, if several provinces get richer, the amount available to the center to transfer to the poor provinces will be larger.

3.3. Transfers, fiscal incentives, and corruption

Studies about the relationship between decentralization and corruption create a puzzle. Cooter (2003), Kotsogiannis and Schwarzger (2006), Shah (1997b), and Shleifer and Vishny (1993) argue that greater decentralization implies less corruption, in part because competition among subnational governments constrains their behavior and policy choice. In contrast, Treisman (2000) argues that federal systems are more corrupt than non-federal ones.

The above comparative theory of federalism provides the answer, showing that not all forms of decentralization are likely to improve welfare. In particular, not all forms of decentralization affect corruption in the same way. Because competition among subnational governments is one of the mechanisms for policing corruption, decentralization must satisfy the conditions of a common market (including mobile factors of production), have sufficient subnational policy authority, and a hard budget constraint (i.e., it must satisfy conditions F2–F4). Most decentralized countries fail to satisfy these conditions; they therefore fail to prevent corruption.

In this subsection, I investigate a second relationship involving how the fiscal system affects corruption, revealing how greater subnational revenue dependence on transfers implies greater corruption. Political officials generate political support in two broad ways – through the provision of market-enhancing public goods and through patronage, rent-creation, and corruption. Both create value for citizens, showing that not all forms of decentralization are likely to improve welfare. In particular, not all forms of decentralization affect corruption in the same way. Because competition among subnational governments is one of the mechanisms for policing corruption, decentralization must satisfy the conditions of a common market (including mobile factors of production), have sufficient subnational policy authority, and a hard budget constraint (i.e., it must satisfy conditions F2–F4). Most decentralized countries fail to satisfy these conditions; they therefore fail to prevent corruption.

In this subsection, I investigate a second relationship involving how the fiscal system affects corruption, revealing how greater subnational revenue dependence on transfers implies greater corruption. Political officials generate political support in two broad ways – through the provision of market-enhancing public goods and through patronage, rent-creation, and corruption. Both create value for individuals and groups, and both induct citizens to support those in power. Providing rent for constituents creates value for them (often at the expense of other local citizens, but sometimes at the expense of citizens in other regions), inducing them to support those in power.

In contrast to corruption and rents, providing market-enhancing public goods has two separate effects. First, it generates support directly through creating value for citizens. Second, because they expand the local economy, market-enhancing public goods increase local revenue, relaxing the budget constraint. A simple comparative static result shows that increasing the portion of a subnational government’s revenue derived from locally generated revenue leads political officials to substitute more market-enhancing public goods for rent-creation and corruption (Careaga and Weingast, 2003). The reason is that greater revenue capture increases the fiscal incentives of political officials to foster market growth. Conversely, the incentives to engage in rent-creation and corruption increase as subnational governments depend more on the central government for revenue and have low local-revenue generating capabilities.

A second aspect of some transfer systems enhances corruption. A common feature of transfer systems is that the central government provides rules and restraints on local government policymaking authority. In many cases, policies are designed in the center with little local discretion. Unfunded mandates are common. The relevance for corruption is that central government policy control impedes the accountability of subnational governments. Centralized control of subnational government policymaking allows local government officials to blame policy failures on the central government whether the latter is responsible or not. Were the central government’s controls really that insidious, or did the local officials simply fail to work around them? When citizens face information problems about the impediments to policy success, local government officials can engage in corruption while blaming the center.

3.4. Fiscal equivalences

Another important idea is the well-known concept of fiscal equivalence, the matching those being taxed with those receiving the benefits.22 This literature shows that a series of incentive problems arise when the political system de-links taxation and spending decisions, causing spending decisions to deviate from efficient levels (see, e.g., Winer and Hettich, 2006). For example, political incentives prevent higher governments typically from providing local public goods efficiently. The incentive problem arises because voters in a locality believe that the tax costs of their programs are spread across all localities. Higher governments often provide local public goods through “universalism” or “something for everyone”; that is, large collections of similar projects to a majority or more of localities (Inman, 1988; Wallis and Weingast, 2005; Weingast, 1979). Universalism characterizes federal rivers and harbors projects throughout American history and, since WWII, of a range of policy benefits, such as sewage treatment plants, federal poverty relief funds, federal highway funds, and most recently, funds for homeland security. Centralized provision of local public goods creates a common pool problem in which voters and representatives seek larger than efficient projects, that is, pork. Universalism leads higher governments to over-provide local public goods and services (Inman, 1988; Poterba and von Hagen, 1999; Weingast et al., 1981; Winer, 1980).23 The same argument holds for a regional government providing local public goods, mutatis mutandis.

4. Further implications of fiscal incentives

Fiscal incentives affect a surprisingly wide variety of policy choices. In this section, I consider several additional applications of the fiscal incentive approach.

22 Olson (1969) introduced the term, and Oates (1972, pp. 33–35) discussed the idea as the “perfect correspondence” (see also Breton, 1998; Wicksell, 1967).

4.1. Fiscal incentives and the political design of markets

One of the most powerful tools for affecting the economic destiny of a country is its regulatory control over markets. This power is inherently political. The principal question concerns why some countries foster thriving markets while others manipulate markets for political purposes – the difference between Shleifer and Vishny's (1998a) “grabbing hand” versus “helping hand.” No general theory exists of these matters.

In this subsection I explore the fiscal incentive approach to demonstrate that the fiscal system provides surprisingly strong incentives affecting political choice of policies with respect to markets, especially for subnational officials. The fiscal incentive approach suggests that government officials are biased toward market policies that generate more revenue within their fiscal system. When they capture revenue based on broad taxes on economic activity, they have incentives to provide market-enhancing public goods and to create new market opportunities as a means of increasing the fiscal proceeds generated by markets. If in contrast they raise revenue by selling monopoly rights, then officials seek to restrict markets.

Both authoritarian and the weak democratic governments in developing countries have strong fiscal incentives to create monopolies (see North et al.'s discussion of the natural state, 2009, ch. 2). Several forces lead to policies fostering monopolies. Developing countries often have little capacity to tax, so selling rights to access markets may be a revenue-generating expedient. These countries also use rent-creation as a means of controlling violence. Because fighting diminishes their rents, rents give individuals and groups with access to violence an incentive to cooperate rather than fight. Finally, governments in these states tend to be insecure, so officials have short time horizons that discount the long term in favor of more immediate payoffs.

Political officials in developing countries often exchange monopolies, privileges and other rights of limited access for revenue and political support. This type of exchange is a time-tested system that dates back 1000s of years. North et al. (2005, chs. 2–3) argue that this exchange represents one of the political foundations of most developing countries today. Limits on entry and competition create rents in markets that can be shared among important elites, firms granted rights in markets, and the government. Maintaining their rents requires that the elites support the government in power, implying that when powerful groups obtain important rights, they support rather than challenge the government. Moreover, the government's insecurity-induced short time horizon implies that it downplays the long-term economic consequences of its policies.

4.2. An illustration: banking in the early United States and modern Mexico

Wallis et al. (1994) apply the fiscal incentive approach to banking in the early United States; Haber (2008) applies it to a comparative study of the early United States and modern Mexico. As the logic in the previous subsection shows, the most natural way for an authoritarian or a weak democratic government in a developing economy to structure the banking industry is to create monopolies; for example, by limiting entry and selling bank charters as a means of creating economic rents that can be shared among the banks, the government, and privileged citizens and firms who receive scarce loans. Because the government has significant interests in banking, exchange of privileged rights often explicitly or implicitly grants the government privileged access to loans.

As Haber (2008) argues, this fiscal-induced industrial organization of the banking sector means that it fails to provide the basic banking functions of an economy, notably, mobilizing capital to highest valued users who create new enterprises or seek to expand profitable ones. Instead, most loans go to the government, insiders, high government officials, and their relatives. An inevitable consequence of this structure is limited competitiveness of the financial sector and constraints on the ability of banks to foster long-term economic growth.

This logic reflects how Mexico has typically structured its banking industry (Haber, 2008). Because charters are valuable, the government has a fiscal interest in restricting competition in the banking industry as a means of increasing its revenue. Haber also shows that Mexican government banking policy has gone through several cycles of rent creation and expropriation of bank assets, a policy cycle all too common in many developing countries.

The early history of the United States yields two important conclusions about fiscal incentives and the political choice of market structure in banking. First, the United States was no exception to the rule about restricting entry to create rents shared among bankers and the government (Wallis et al., 1994). In 1800, most states used this system, including Pennsylvania whose commercial center of Philadelphia was the country's banking center.

Second, the United States had a strong market-preserving federal structure throughout the late eighteenth and nineteenth centuries, which affected states' fiscal interest. Under condition F2, states had nearly exclusive regulatory control over markets within their borders; under F3, states participated in a common market with product and factor mobility; and, under F4, they faced a hard budget constraint. Moreover, states raised virtually all of their own revenue. This structure allowed states the freedom to design and redesign the rules governing various markets.

In the decade following 1800, Massachusetts slowly switched systems. Beginning with the monopoly approach, it created one large bank in which it invested heavily and several smaller banks. The state also imposed a tax on bank capital, which worked against the smaller banks: as the majority owner of the large bank, the state effectively paid part of its own tax. Yet over time, the state found it raised more revenue from taxes on the smaller banks than it did in dividends from the large bank. The state's fiscal incentives led it to alter the regulatory system: it sold its interest in the large bank, and it stopped limiting entry and selling charters, and instead combined relatively low taxes on bank capital with more open entry into banking. The new system gave Massachusetts banks a competitive advantage over all other U.S. banks. Merchants, enterprises, and transactions funded in Boston – such as financing, insuring, marketing, and transporting U.S. export crops bound for Europe – had an economic edge over their competitors from other states.

Because a competitive banking sector maximizes the size of its tax base, Massachusetts's fiscal incentives – in contrast to that in all other states, including Pennsylvania – led it to promote a competitive banking sector. This system was so successful that, by the early 1830s, Massachusetts had more banks and more bank capital than any state in the country. It also received over 50 percent of its revenues from the tax on bank capital allowing it to make great reductions in the principal tax falling on its citizens, the property tax. This was a win–win policy for that state.

The competitive system led Massachusetts to eclipse Philadelphia as the nation's banking center. A number of years later, New York also switched fiscal systems, emulating Massachusetts, and New York City eclipsed both Boston and Philadelphia as the nation's banking center. Many other states subsequently switched to the system that worked. Had the United States been a centralized federalism, as modern Mexico, the national government would have had little incentive to alter the original system of limited entry once it was in place. In contrast to competitive among states, the lack of competition facing the national government would not
have provided incentives for it to innovate and foster a competitive banking system.

4.3. Other illustrations

(1) Shleifer and Vishny’s (1998b) study of the differential local government support for the economy in Poland and Russia draws on the fiscal incentive logic. Using survey techniques, they find that local governments in Poland are far more supportive of business than in Russia. Shleifer and Vishny attribute this difference to government officials’ incentives. First, “the [electoral] incentives of local politicians in Russia – unlike those in Poland – do not encourage them to support private business” (Shleifer and Vishny, 1998b, p. 248). Second, local government fiscal interests differ significantly. In Poland, local governments rely on local taxes, fees, and property taxes, so fostering local economic prosperity yields greater revenue. In Russia, most local revenue comes from higher governments that exhibit considerable opportunism with respect to transfers to lower government: regional governments reduce their transfers to cities that increase their revenue. Fostering a healthier local economy by Russian cities fails to generate greater revenue for the local government. Consistent with Zhuravskaya’s (2000) findings noted above, local officials have very low fiscal incentives to foster economic growth. As Shleifer and Vishny (1998b, p. 249) conclude, “The effects of such fiscal federalism... are perverse.”

This discussion yields three general lessons. First, consistent with the theme of this section, a government’s fiscal incentives has strong effects on its incentives to choose pro- or anti-market policies. Governments that raise money from broad and relatively uniform taxes on general economic activities are far more likely to choose policies that foster markets. Governments that raise revenue through restrictive economic activities instead manipulate markets for political and fiscal ends. Second, subnational governments’ fiscal incentives in the presence of inter-jurisdictional competition is one of the ways a market-preserving federal system promotes pro-market policies. Third, opportunistic higher governments that capture the revenue gains from increased local economic activity reduce incentives for subnational governments to foster economic growth.

(2) The familiar logic soft budget constraints is also a fiscal incentive story. Fiscal problems associated with soft budget constraints have led scholars to study the incentive effects of the hard and soft budgets for subnational governments (see, e.g., Dillinger and Webb, 1999; Haggard and Webb, 2004; Kornai, 1986; McKinnon, 1997; Rodden et al., 2001; SanguINETTI, 1994; Wibbels, 2003). A soft budget constraint arises when subnational governments believe they will be able to externalize some of their fiscal burdens, such as a national bailout subnational governments in fiscal distress. Subnational governments facing a soft budget constraint have a reduced (or no) fiscal incentive to make prudent financial decisions.

(3) Timmons (2005) provides evidence for the incentives interest model. He shows that voluntary compliance with taxes is higher when citizens received public goods and services they value in exchange for taxes (see also Levi, 1988). This effect is especially strong among subnational governments, particularly in a market-preserving federal context.

5. Democracy and decentralized governance

Democracy is perhaps the most celebrated institutional means of creating political accountability. Perhaps most ostensibly, elections potentially allow citizens to influence their own destiny by choosing among potential officials. Voting also helps citizens to help police their rights. As James Madison emphasized in the Federalist Papers, voting allows citizens to “throw the rascals out” (see Riker, 1982). The threat of being turned out of office provides political officials with incentives to make decisions that reflect their constituents’ interests, including honoring citizen rights.24

This section discusses the interaction of decentralization and democratic governance. It suggests how fiscal decentralization strengthens democracy. Before we turn to this interaction, however, we must understand some of the limits of democracy, elections in particular.

5.1. Potential limits of democracy

Democracy is so attractive that policymakers and donor organizations too often fail to study the conditions under which it is more likely to succeed. Three aspects of democracy are critical for our analysis of democracy, one empirical and two theoretical.

The empirical aspect is that most new democracies fail, typically due to coups or to “democratic set-asides” (incumbents cancel elections or refuse to step down after losing an election). The evidence is striking that democracy is far more likely to succeed in richer countries. Przeworski (2006) estimates that the frequency of a democracy failing each year in a country with a per capita income of less than $1000 per year is 0.085 or one in twelve (see Table 2); with a per capita income of $3001–6055, it is 0.016 or one in sixty-one; while no democracy with a per capita income of greater than $6055 has failed.25 Put another way, a democracy in the poorest category has only a 0.41 chance of remaining a democracy one decade later; in the $1000–3000 category, the probability is 0.69; in $3001–6055, 0.85, and 1.00 for the highest income category.

The first theoretical aspect of democracy are the costs and dangers it can impose on citizens. Elections empower governments to tax, regulate business, define property rights, and jail people. All these powers can be abused, as tyranny of the majority suggests; and even if not abused, these powers may impose sufficiently large costs that some citizens support extra-constitutional action and violence as a means of defending themselves.

The dangers of democracy are difficult for people in the developed West to understand because democracy in these countries allows citizens to determine their own destiny. But democracy in these countries is embedded in a series of institutions and norms that complement elections by place striking limits on government policymaking and which protect citizens from many potential abuses. Courts and other institutions, for example, enforce a wide range of citizen rights; and elaborate procedures constrain the range of feasible policies. Yet democracy in the developing

<table>
<thead>
<tr>
<th>Income levela</th>
<th>Estimated probability of failure per year</th>
<th>Estimated probability of surviving 10 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; $1000</td>
<td>0.085</td>
<td>0.41</td>
</tr>
<tr>
<td>1000–3000</td>
<td>0.036</td>
<td>0.69</td>
</tr>
<tr>
<td>3001–6055</td>
<td>0.016</td>
<td>0.85</td>
</tr>
<tr>
<td>&gt; 6055</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>


24 Riker (1982) provides a systematic analysis of these two aspects of democracy, emphasizing the importance of the second. Besley (2006) and Persson and Tabellini (2000) provide comprehensive analyses of incentives, accountability, and responsiveness generated by electoral and political institutions.

25 Przeworski’s figures are in 1985 purchasing parity dollars. Moreover, for several reasons, these estimates should be taken as indicative. The data are necessarily derived from post-WWII history, so there is no event like the Great Depression in whose wake many democracies failed.
context typically lacks these complementary institutions that help sustain it.

The second theoretical aspect of democracy concerns the limit condition; all successful democracies limit the stakes of power through various institutions that restricting the scope of policy authority of elected representatives (Weingast, 2009). Successful democracies limit the stakes of power protect a range of citizen rights and other aspects of the status quo. The centrality of the limit condition is revealed by events in Chile in 1973 where the legitimately elected government threatened landowners and others on the political right, leading them to support a bloody coup. When citizens believe they are protected under the system, they are far less likely to support extra-constitutional action, such as coups. Democracies that satisfy the limit condition are therefore more stable.

The absence of the limit condition in the developing context reveals a critical difficulty with sustaining democracy in the poorest and under-institutionalized countries. These states face grave difficulties maintaining institutions that satisfy the limit condition.

5.2. Democracy and decentralization

Decentralization has a surprisingly wide range of interactions with democratic governance. First, fiscal decentralization contributes to the limit condition in two different ways. Under market-preserving federalism, local public goods provision – such as water, electricity, fire protection, and garbage – is typically independent of who holds the national government. Surprisingly, in the developing context, centralized governments often use their fiscal discretion to cut off funds to opposition areas (see the discussion of tragic brilliance in Section 5.3). Because decentralization makes subnational government resource allocation independent of national elections, decentralization lowers the stakes of losing the national government.

Decentralization also contributes to the limit condition through party politics. In a centralized nation, losing a national election is very costly to the incumbent and its supporters. Incumbents therefore have incentives to hold onto power despite losing. In the absence of the limit condition, it may simply be too costly to give up power. In a decentralized state, however, losers can typically maintain a local power base from which to remain politically visible and to provide some benefits to their constituents. Local political strongholds also provide a base from which this party can launch a future attempt to recapture national power. By lowering the stakes of power, decentralization makes it more likely that losers of national elections will give up power.

Second, a growing literature studies the relationship between decentralization and conflict in divided societies; that is, states with ethnic, cultural, or linguistic differences. In some cases, decentralizing authority to regions with more homogeneous populations allows these groups to live in harmony within a larger state, which seems to play a role in “holding together,” including Belgium, India, Spain, and the Netherlands (Lijphart, 1975; Stepan, 2004a). Decentralization has also seemed to mitigate conflict in the Indonesia and the Philippines. Inman and Rubinfeld (2008) argue that decentralization was essential to the democratic transition in South Africa, both helping to limit expropriation and to make democracy self-enforcing. They show that whites were willing to support the transition to democracy because decentralization provided them with sufficient security.

In contrast, decentralization sometimes exacerbates conflict in divided societies (Snyder, 2000). Eaton (forthcoming), for example, argues that decentralization in Colombia exacerbated its conflict because the control of local governments provide the different groups with resources and authority useful for fighting. The bottom line is that we know too little about whether decentralization – or specific forms of decentralization – help mitigate conflict.

Third, Myerson (2006) argues that decentralization adds to the success of democracy in another way; namely, to help incubate candidates for national office. Subnational office allows officials to gain experience and reputation. Decentralization therefore provides national voters with more information about the candidates, and they can pick for national leaders those who have been especially successful at the subnational level.

Finally, a large literature suggests that certain institutional features of democracy are more likely to preserve decentralization. I have already mentioned Riker’s thesis about the party system in this context (see Section 2). The literature associates a range of institutions with stable decentralization; for example, when subnational officials are elected, in contrast to serving at the pleasure of the national government; or when the constitution designates that the principal subnational units have direct representation in the government (e.g., in a “senate”).

5.3. Tragic brilliance: how insecure governments use centralized fiscal control to undermine elections

Elections create a problem in the developing context sometimes called the “tragic brilliance” mechanism by which a centralized authoritarian or weak democratic regime perverts elections so they serve as a mechanism of social control rather than citizen choice (Diaz-Cayeros et al., 2008). As noted, democracy in the developed West satisfies the limit condition – these countries impose credible limits on what democratically elected representatives may do (Weingast, 2009). Citizens enjoy a wide range of rights and public goods and services by virtue of citizenship, not based on a political relationship with those in power. In particular, standard local public services – such as water, electricity, education, sewer service, and road maintenance – do not depend on whom an individual or a locality votes for.

Elections in many authoritarian and weak democratic regimes often differ dramatically from this ideal. In Mexico under the PRI – the party that virtually monopolized power from 1930 through the early 1990s – elections served a very different purpose than citizen choice. Although Mexico has long been a federal system, the PRI engineered a very centralized one (Diaz-Cayeros, 2006), where the central government raises most of the revenue and finances most state and local expenditures through transfers. In the 1980s, the average local government received over 80 percent of its revenue from higher governments.

Although this pattern of revenue generation and transfer conforms to that recommended by FGFF, its purpose was not to further citizen welfare. Instead, the PRI used its centralized discretion over revenue to threaten localities who supported the opposition by withdrawing funds to finance local governments. The threat to withdraw revenue forced opposition-favoring citizens to face a dilemma: voting for the opposition meant a far smaller level of public services. Revenue centralization afforded the PRI the discretion to force most voters to support it at the polls, even voters who preferred the opposition.

As evidence, the case study literature shows that when the first two cities, Ciudad Juarez and Chihuahua, voted in the opposition in 1983, the cities lost on the order of half their budgets (Rodriguez, 1995; Rodriguez and Ward, 1995). Similarly, in a study of

---

26 Siegle (2006) surveys this literature.

27 Bland (2006) surveys this literature. See also the references in note 7, infra, especially Dillinger and Webb (1999) and Carman et al. (2000).

28 Some of the transfers were by formula, but a large portion of it was discretionary, especially for local governments (Careaga and Weingast, 2003).
1800 of 2400 Mexican municipalities from a more recent period, Diaz-Cayeros et al. (2008) provide econometric evidence showing that municipalities which support the opposition receive on average one-quarter less revenue.

The tragic brilliant mechanism represents a pathology of both democracy and decentralization. It is tragic because it forces opposition-leaning citizens to play an active role in maintaining a regime that they would rather replace; but also brilliant in that authoritarian leaders use their policy discretion to create political dependence and subservience while providing the outward veneer of choice and democracy.

The tragic brilliance mechanism reveals a political motivation for why regimes in developing countries centralize policy authority and taxation. Wholly apart from administrative efficiencies and fiscal equalization, centralization affords regimes with political leverage over lower governments and citizens. By making the delivery of basic local public goods and services depend on whom citizens vote, the incumbent regime compromises citizens' ability to throw the rascals out, to exercise fiscal autonomy, and to influence public policies.

The main lesson is that, for democracy to serve as a mechanism of freedom and choice, it must be embedded in institutions that constrain the government's use of discretionary fiscal authority to threaten voters who vote for the opposition. Fiscal decentralization that satisfies condition F5 lowers the ability of those in power to exercise tragic brilliance. Preventing the tragic brilliance mechanism therefore provides another SGFF rationale for decentralizing fiscal authority. Independent taxation authority allows local governments not only a fiscal interest in fostering local economic prosperity, but also a much greater degree of independence from a controlling (and potentially predatory) center.

6. Overcoming impediments to decentralization

In this section, I consider two strategies for implementing decentralization, especially in the presence of predatory governments.

6.1. Local government fiscal independence in predatory systems

Predatory central governments are a problem throughout the developing world, and these governments can hinder the operation of an otherwise well-designed federal system. A predatory central government that faces relatively few constraints on its behavior can reverse or compromise any and all of the benefits of decentralization.

6.1.1. The problem

A common and yet insidious form of predation perverts the logic of innovation and competition in a local government exhibiting policy independence. Suppose a particular subnational government creates a thriving local economy whose performance that stands out in comparison with other regions. This economic success potentially provides local political officials with a resource and political base with which to challenge national leaders, either to extract greater concessions or freedoms; or to challenge their leadership. The threat of political challenge provides predatory or insecure central governments with an incentive to prevent local governments from succeeding. This threat suggests that national leaders of predatory states have an incentive to use their powers to reduce or remove the authority of the local government; they can expropriate control of all successful enterprises; or they can take over the local government and reverse its policies. Crook and Manor (1998) provide an instructive example of how the dominant Congress party in India dismantled the opposition Janata Party's successful ruling of the state of Karnataka in the 1980s.

The economic side of this problem is even worse. The risk of a predatory reaction by the central government feeds back into the local economy, making it less likely that economic agents will make investments that can be expropriated even if these would be profitable under the local government's policies. It may well be that, in truly predatory governments, there is little hope for reform. In combination, these two effects diminish subnational government officials incentives to foster economic performance.

6.1.2. Overcoming predation through decentralization

Under some circumstances, a formerly predatory government may seek reform. How is it to improve economic performance? China's successful creation of market-preserving federalism suggests one way around these problems. The last condition of market-preserving federalism requires institutional limits provide some form of credible commitment by the central government to honor the rules of the federal system. Whether by design of happenstance, China's reform-minded leaders accomplished this condition in two ways. The first and perhaps more important is fiscal.

Communist China had a long history of anti-market policies, mass murder, and other forms of predation. This predatory behavior represented a strong impediment to market reform, economic growth, and investment: why should economic agents trust such a government to honor economic reform policies rather than, at some point down the road, reverse itself and punish those successful under reform? Communist China under Mao exhibited several great policy reversals with exactly that type of punishment; notably, the Great Leap Forward and the Cultural Revolution.

This political risk meant that economic reform had, somehow, to limit the authority of the central government. China's strategy in promoting economic reform was decentralization: the devolution of economic policymaking (policy authority condition) and fiscal authority (including condition F3, the HBC) to the provinces (Montinola et al., 1995; Oi, 1992; Shirk, 1993). This new policy authority allowed the reform-oriented provinces to alter their policies from socialist to pro-market. Provinces also faced strong fiscal incentives to promote reform under what the Chinese called the "fiscal contracting system," 1981–1992 (see Jin et al., 2005; Oksenberg and Tong, 1991). Under this system, most provinces raised their own taxes under a fiscal contract with the central government. Many of the fiscal contracts were step functions: share fifty percent of all revenue raised up to some level and then allow the province to retain 100 percent of all revenue beyond. The average province faced a marginal tax retention rate of 89 percent (Jin et al., 2005). Reflecting strong fiscal incentives to promote reform, many provinces quickly grew rich as their economies mushroomed.

In keeping with SGFF logic, as the reforms succeeded, the combination of policy authority and fiscal health granted the provinces both the incentive and political power to act independently of the central government. The fiscal incentives also had strong political effects on constraining the central government. Because most provinces benefited from and had significant investment in the system, they were able to counterbalance the central government. Consider, for example, the conservative reaction against reform followed the suppression of the protests in Tiananmen Square in 1989. Provincial leaders – in particular, the governor of the most successful reform province, Guangdong – used their independence to prevent the proposed anti-reform reaction (Shirk, 1993, pp. 194–195; Montinola et al., 1995).

A second mechanism arose to raise the costs to the central government of an anti-market reaction, although this one was not by design. An important aspect of Chinese economic reform is the floating labor population, workers from the interior who come to the coastal reform provinces to work (Solinger, 1999). These laborers cannot become local citizens but instead work under a system...
of limited rights – effectively an intra-China guest worker system. Host provinces retain the right to send these laborers back to their home provinces. This labor population is now huge – over 100 million workers.

The existence of a huge floating labor system has a striking political implication: the most likely response to an anti-reform reaction by the central government would be for the reform provinces to kick out many or most of the floating laborers. This means that 10s of millions of people – perhaps approaching 100 million – would instantly become problems for the central government: how would they be fed, clothed, and housed? Because hungry people can topple governments, this potential reaction places a significant hurdle in front of political leaders who are tempted to impose an anti-market reaction to economic reform.

The main implication is straightforward: although China’s devolution of power to the provinces at the inception of economic reform was under the discretion of the central government, the reform’s success created strong power centers in the provinces that counterbalanced the center’s discretion.

### 6.2. Decentralizing one step ahead

The Chinese case also suggests an important strategy for implementing decentralization. Many developing countries face resistance to decentralization. Indeed, poorly designed decentralization has made things worse in some countries, for example, due to soft budget constraints or to mismatches in responsibility and resources. In many developing countries, an across-the-board decentralization may therefore be problematic. The political and economic situation of some localities is such that greater freedom will result not in greater responsiveness to local citizen welfare, but instead greater authority and resources allow local officials to create a larger scope for the system of local rents and corruption (Haggard and Webb, 2004).

An alternative to across-the-board decentralization is to decentralize in a series of steps. The idea is first to identify a province or region that is most likely to succeed in fostering local economic growth; and then to design decentralization so that this province obtains new authority, incentives, and resources to reform “one step ahead.” The purpose of this strategy is to create a demonstration effect that decentralization can work in this country.

The Chinese successfully employed this strategy, allowing Guangdong Province to begin their reform process before other provinces gained that authority. Many other provinces were skeptical of reform, and used their increased powers to maintain or even reinforce the traditional system. But Guangdong’s quick success won converts around the country, and even some of the most traditional provinces embraced reform. For example, Heilongjiang Province, ideologically committed to the communist regime, reacted to Guangdong’s market reforms by increasing the standard subsidies of the socialist system. Yet Guangdong’s reforms succeeded in lowering market prices of the same goods below the subsidized prices elsewhere – and without the fiscal costs of subsidies. Heilongjiang’s fiscal incentives then led them to dismantle their expensive subsidies and imitate Guangdong (see Montinola et al., 1995).

A similar, one step ahead strategy has emerged in Mexico in the areas seeking to integrate with the United States economy, and to a lesser degree, in India. In Mexico, the center actively discouraged this independent movement from below (per discussion of the tragic brilliance mechanism above in Section 5.3.), but could not prevent it. Many of the export-localities wrestled political control from the dominant party, the PRI, in order to improve the delivery of local services necessary to foster the light export industry developing in Northern Mexico. Although the central authorities punished these areas with a marked decline in revenue transfers, the localities made up the revenue deficit by removing corruption – the PRI used their control of local utilities to pad the labor budget by mailing money to supporters throughout Mexico – and by charging user fees for improved local services (Rodriguez, 1995; Rodriguez and Ward, 1995). As Rodriguez (1995, p. 166) suggests, “Over the course of only a few years, the ratio of state to local revenues... changed from around 70 percent state funding to over 70 percent local funding.” Citizens and firms willingly paid user fees for reliable, valued services, such as solid waste disposal, water, and road maintenance.

Part of the reason this system works is the high demand for more efficient local public goods and services necessary to integrate the economy with the United States. The success of the first two municipalities to attempt this strategy, Ciudad Juarez and Chihuahua in 1983, created the demonstration effect. By the mid-1990s, most of the larger cities in Mexico were governed by the opposition.

### 7. Conclusions

FGFF and SGFF approaches are complementary rather than competing. FGFF studies the optimal design of fiscal institutions in the context of welfare maximization without respect to the incentives of political officials. SGFF extends and adapts FGFF lessons to the context of incentives and self-interested political officials.

One difference is that SGFF attempts to make explicit the political assumptions underlying FGFF’s prescriptions, as represented by the conditions of market-preserving federalism. This perspective also helps identify the incentives facing political officials under different forms of decentralization. FGFF scholars have always understood the importance of policy authority, common market and HBC for their prescriptions. Because these conditions were often implicit in the FGFF framework, many decentralizations in the last 20 years have been designed without attention to these conditions. Making them explicit also makes clearer the pathologies arising when decentralization fails to satisfy one or more of these conditions.

The SGFF approach also amends many FGFF normative lessons, for example, those concerning intergovernmental transfer systems. FGFF tends not to study the incentive effects of transfer systems, and many transfer systems around the world provide political officials with poor incentives to foster local economic prosperity. SGFF provides several lessons for the design of transfer systems. First, it emphasizes the critical importance of local government revenue generation. Local revenue generation makes local governments more responsive to citizens, reduces corruption, and increases the incentives to provide market-enhancing public goods. Local revenue generation is also important in a political sense. Central governments of many developing countries decentralize, but with too many strings and conditions that compromise the effects of decentralization. Subnational governments often have few resources with which to resist the center. Fiscal independence grants subnational governments bargaining leverage and hence a degree of political independence. Second, SGFF emphasizes the importance of step functions in transfer systems to provide subnational governments with higher marginal incentives to foster local economic prosperity.

More generally, the fiscal incentives approach shows that the form of the tax system affects subnational government policymaking, particularly policies with respect to the market. All governments have a bias toward policies that increase their revenue. Because market-enhancing public goods increase their tax revenue, governments that rely on broad-based taxes are more likely to foster local economic prosperity than governments that rely on privileges and monopolies for their revenue. SGFF approaches also emphasize that greater marginal revenue retention by local gov-
ernments increases incentives of local political officials to provide market-enhancing public goods. Greater marginal tax retention increases the fiscal return from these goods and therefore makes them more attractive to local public decisionmakers.

Perhaps the greatest area of blending of FGFF and SGFF approaches is with respect to financial mechanisms. All scholars now recognize the critical importance of establishing hard budget constraints for all levels of government, especially local ones. Soft budget constraints give poor incentives and lead to a range of financial and economic problems.

Finally, SGFF logic also suggests that fiscal decentralization interacts with democracy. First, decentralization contributes to the limit condition, lowering the stakes of national elections and therefore making democracy more stable. Second, elections potentially increase accountability. But democracy can only play this role in the absence of the tragic brilliance mechanism. Fiscal decentralization inhibits tragic brilliance by limiting the ability of governments in developing countries to threaten citizens by withdrawing local public goods if they support the opposition.

Acknowledgments

The author thanks Robert Aten, Jan Brueckner, Edwin Connerley, Jurgen von Hagen, Robert Inman, Jose Larios, Gary Libecap, Charles McLure, Roger Noll, Wallace Oates, Yingyi Qian, Jonathan Rodden, Daniel Rubinfeld, Kenneth Shepsle, Nirvikar Singh, Ernesto Stein, Jamie Thomson, Craig Volden, and John Walls for helpful conversations. This paper was commissioned by ARD, Inc. under a contract with the Office of Democracy and Governance of the U.S. Agency for International Development. The paper was designed to inform the Agency’s revision of its “Decentralization and Democratic Local Governance Programming Handbook.”

References


Enikolopov, Ruben, Zhuravskaya, Ekaterina, 2002. Influence of political institutions on the effect of decentralization Center for Economic and Financial Research (CEFIR), Moscow, Russia.


Shah, Anwar, 1997a. The quest for the right balance and responsive and accountable governance: Lesson from decentralization experience. WP, WB.


