Inside

The New Structural Economics
Comments on The New Structural Economics
Gender and Firm Creation
Enterprise Performance
Business Environments and Development

Volume 26 • Number 2 • August 2011

www.wbro.oxfordjournals.org
The World Bank Research Observer

EDITOR
Emmanuel Jimenez, World Bank

CO-EDITOR
Luis Servén, World Bank

EDITORIAL BOARD
Harold Alderman, World Bank
Barry Eichengreen, University of California-Berkeley
Marianne Fay, World Bank
Jeffrey S. Hammer, Princeton University
Ravi Kanbur, Cornell University
Howard Pack, University of Pennsylvania
Ana L. Revenga, World Bank
Ann E. Harrison, World Bank

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.

# Research Observer

**Volume 26 • Number 2 • August 2011**

## Symposium on “New Structural Economics”

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Structural Economics: A Framework for Rethinking Development</td>
<td>193</td>
</tr>
<tr>
<td>Justin Yifu Lin</td>
<td></td>
</tr>
<tr>
<td>Comments on “New Structural Economics” by Justin Yifu Lin</td>
<td>222</td>
</tr>
<tr>
<td>Anne Krueger</td>
<td></td>
</tr>
<tr>
<td>Comments on “New Structural Economics” by Justin Yifu Lin</td>
<td>227</td>
</tr>
<tr>
<td>Dani Rodrik</td>
<td></td>
</tr>
<tr>
<td>Rethinking Development Economics</td>
<td>230</td>
</tr>
<tr>
<td>Joseph E. Stiglitz</td>
<td></td>
</tr>
</tbody>
</table>

## Symposium on “Business Environments”

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender and the Business Environment for New Firm Creation</td>
<td>237</td>
</tr>
<tr>
<td>Leora F. Klapper and Simon C. Parker</td>
<td></td>
</tr>
<tr>
<td>Explaining Enterprise Performance in Developing Countries with Business Climate Survey Data</td>
<td>258</td>
</tr>
<tr>
<td>Jean-Jacques Dethier, Maximilian Hirn, and Stéphane Straub</td>
<td></td>
</tr>
<tr>
<td>The Effects of Business Environments on Development: Surveying New Firm-level Evidence</td>
<td>310</td>
</tr>
<tr>
<td>Lixin Colin Xu</td>
<td></td>
</tr>
</tbody>
</table>
As strategies for achieving sustainable growth in developing countries are re-examined in light of the financial crisis, it is critical to take into account structural change and its corollary, industrial upgrading. Economic literature has devoted a great deal of attention to the analysis of technological innovation, but not enough to these equally important issues. The new structural economics outlined in this paper suggests a framework to complement previous approaches in the search for sustainable growth strategies. It takes the following into consideration.

First, an economy’s structure of factor endowments evolves from one level of development to another. Therefore, the optimal industrial structure of a given economy will be different at different levels of development. Each industrial structure requires corresponding infrastructure (both “hard” and “soft”) to facilitate its operations and transactions.

Second, each level of economic development is a point along the continuum from a low-income agrarian economy to a high-income industrialized economy, not a dichotomy of two economic development levels (“poor” versus “rich” or “developing” versus “industrialized”). Industrial upgrading and infrastructure improvement targets in developing countries should not necessarily draw from those that exist in high-income countries.

Third, at each given level of development, the market is the basic mechanism for effective resource allocation. However, economic development as a dynamic process requires industrial upgrading and corresponding improvements in “hard” and “soft” infrastructure at each level. Such upgrading entails large externalities to firms’ transaction costs and returns to capital investment. Thus, in addition to an effective market mechanism, the government should play an active role in facilitating industrial upgrading and infrastructure improvements. JEL codes: L16, O10, O20, O21, O25, O40
Several decades from now when economic historians look back on the story of the past hundred years, it is very likely that they will be intrigued by the mystery of diverging performances by various countries, especially during the second half of the twentieth century. On the one hand, they will be amazed by the rapid growth path followed by a small number of countries such as Brazil, Chile, China, Indonesia, India, Korea, Malaysia, Mauritius, Singapore, Thailand, and Vietnam, where the industrialization process quickly transformed their subsistence, agrarian economies and lifted several hundred million people out of poverty in the space of one generation. On the other hand, they will be puzzled by the apparent inability of many other countries, where more than one-sixth of humanity remained trapped in poverty. They will also notice that with the exception of a few successful economies, there was little economic convergence between rich and poor countries in spite of the many efforts made by developing countries and despite the assistance of many multilateral development agencies.

Long-term sustainable and inclusive growth is the driving force for poverty reduction in developing countries, and for convergence with developed economies. The current global crisis, the most serious one since the Great Depression, calls for a rethinking of economic theories. It is therefore a good time for economists to reexamine development theories as well. This paper discusses the evolution of development thinking since the end of World War II and suggests a framework to enable developing countries to achieve sustainable growth, eliminate poverty, and narrow the income gap with the developed countries. The proposed framework, called a neoclassical approach to structure and change in the process of economic development, or new structural economics, is based on the following ideas:

- **First**, an economy’s structure of factor endowments evolves from one level of development to another. Therefore, the industrial structure of a given economy will be different at different levels of development. Each industrial structure requires corresponding infrastructure (both tangible and intangible) to facilitate its operations and transactions.
- **Second**, each level of economic development is a point along the continuum from a low-income agrarian economy to a high-income post-industrialized economy, not a dichotomy of two economic development levels (“poor” versus “rich” or “developing” versus “industrialized”). Industrial upgrading and infrastructure improvement targets in developing countries should not necessarily draw from those that exist in high-income countries.
- **Third**, at each given level of development, the market is the basic mechanism for effective resource allocation. However, economic development as a dynamic process entails structural changes, involving industrial upgrading...
and corresponding improvements in “hard” (tangible) and “soft” (intangible) infrastructure at each level. Such upgrading and improvements require an inherent coordination with large externalities to firms’ transaction costs and returns to capital investment. Thus, in addition to an effective market mechanism, the government should play an active role in facilitating structural changes.

The remainder of the paper is organized as follows: the next section examines the evolution of development thinking and offers a critical review of some of its main schools of thought. I then outline the basic principles and conceptual framework of the new structural economics, the function of the market, and the roles of a facilitating state. In the next section I highlight similarities and differences between old and new structural economics, and discusses some preliminary insights on major policy issues based on this new approach.

A Short Review of Development Thinking and Experiences

The process of sustainable per capita income increase and economic growth, characterized by continuous technological innovation and industrial upgrading, is a modern phenomenon. From Adam Smith to the early twentieth century, most economists believed that laissez-faire was the best vehicle for achieving sustainable growth in an economy. It was assumed that in thriving economies all decisions about resource allocation are made by economic agents interacting in markets free of government intervention. The price system determines not only what is produced and how but also for whom. Households and firms pursuing their own interests would be led, “as if by an invisible hand,” to do things that are in the interests of others and of society as a whole. Although the laissez-faire approach was challenged by Marxist economists and others, it became the dominant intellectual framework for the study of growth in all countries and remained so for a long time. It certainly provided many good insights on the process of economic development but it missed the importance of the process of continuous, fundamental technological changes and industrial upgrading, which distinguishes modern economic growth from premodern economic growth (Kuznets 1966).

The study of economic development proceeds in two related but separate tracks: growth theories and development theories. While some of the key ingredients of modern growth theory such as competitive behavior, equilibrium dynamics, the importance of physical capital and human capital, the possibility of diminishing returns, and the impact of technological progress can be found in the work of classical economists (Ramsey 1928; Schumpeter 1934), systematic
modeling only started in the 1940s when some pioneers used primary factors to build generic models based on aggregate production functions. Harrod (1939) and Domar (1946) triggered extensive research along these lines. Following their initial work, the Solow-Swan model sparked the first major wave of systematic growth analysis. The objective was to understand the mechanics of growth, identify its determinants, and develop techniques of growth accounting, which would help explain changes in the momentum and role of economic policy. That first generation of growth researchers highlighted the centrality of capital. One important prediction from these models was the idea of conditional convergence, derived from the assumption of diminishing returns to capital—poor economies with lower capital per worker (relative to their long-run or steady-state capital per worker) will grow faster. While that assumption allowed the model to maintain its key prediction of conditional convergence, it also seemed odd: technology, the main determinant of long-run growth, was kept outside of the model (Lin and Monga 2010).

A new wave of growth modeling had to come up with a convincing theory of technological change. Endogenous growth theory, as it came to be known, maintained the assumption of nonrivalry because technology is indeed a very different type of factor from capital and labor—it can be used indefinitely by others, at zero marginal cost (Romer 1987, 1990; Aghion and Howitt 1992). But it was important to take the next logical step and to understand better the public good characterization of technology and think of it as a partially excludable nonrival good. The new wave therefore reclassified technology not just as a public good but as a good that is subject to a certain level of private control. However, making it a partially excludable nonrival good and therefore giving it some degree of excludability or appropriability was not sufficient to ensure that incentives for its production and use were socially optimal. The move away from perfect competition was therefore necessary. It has yielded high methodological payoffs. While neoclassical models of growth took technology and factor accumulation as exogenous, endogenous growth models explain why technology grows over time through new ideas and provide the microeconomic underpinnings for models of the technological frontier.

Another important question has been to understand how technological diffusion takes place across countries and generates or sustains growth—and why it does not take root in others. Various interesting possibilities have recently been explored in an attempt to answer that critical question (Jones 1998; Acemoglu, Johnson, and Robinson 2001; Glaeser and Shleifer 2002). Both on the theoretical and empirical fronts, progress has been made in our understanding of growth in recent decades. However, growth research still faces significant methodological difficulties and challenges in identifying actionable policy levers to sustain and accelerate growth in specific countries.
Intellectual progress has been even slower in the particular domain of development theories. It took a paper by Rosenstein-Rodan (1943) to bring development issues to the forefront of the discipline. The paper suggested that the virtuous circle of development depended essentially on the interaction between economies of scale at the level of individual firms and the size of the market. Specifically, it assumed that modern methods of production can be made more productive than traditional ones only if the market is large enough for their productivity edge to compensate for the necessity of paying higher wages. But the size of the market itself depends on the extent to which these modern techniques are adopted. Therefore, if the modernization process can be started on a very large scale, then the process of economic development will be self-reinforcing and self-sustaining. If not, countries will be indefinitely trapped in poverty.

Rosenstein-Rodan’s framework sparked a wave of similar ideas (Chang 1949; Lewis 1954; Myrdal 1957; Hirschman 1958) which came to be known as the structuralist approach to economic development. These early development theories held that the market encompassed insurmountable defects and that the state was a powerful supplementary means to accelerate the pace of economic development (Rosenstein-Rodan 1943; Nurkse 1953; Hirschman 1958). The slump of international trade in the Great Depression led to export pessimism in the post-War period. In Latin America, for instance, political leaders and social elites were influenced strongly by the deterioration in the terms of trade, the economic difficulty encountered during the Great Depression in the 1930s, and the thesis developed by Prebisch (1950) and Singer (1950). They believed that the decline in the terms of trade against the export of primary commodities was secular and led to the transfer of income from resource-intensive developing countries to capital-intensive developed countries. They argued that the way for a developing country to avoid being exploited by developed countries was to develop domestic manufacturing industries through a process known as import substitution. Moreover, the emergence of previous colonies or semi-colonies as newly independent states in Asia and the Middle East, and later in Africa, was accompanied by strong nationalist sentiments.

The results were disappointing in many cases. In many developing countries, well-intended government interventions failed. This was the case across Latin American, African, and South Asian countries in the 1960s and 1970s when import substitution and protection were essential features of the development strategy. One of the main reasons for the failure of many former socialist and developing countries to achieve dynamic growth in their transitional processes was the fact that they attempted to defy the comparative advantage determined by their endowment structures and gave priority to development of capital-intensive heavy industries when capital in their economies was scarce. In order to implement such strategies, developing-country governments had to protect...
numerous nonviable enterprises in their priority sectors (Lin 2009a; Lin and Li 2009).

By shielding unsustainable industries from import competition, developing countries also imposed various types of other costs on their economies. Protection typically led to: (i) an increase in the price of imports and import-substituting goods relative to the world price and distortions in incentives, pushing the economy to consume the wrong mix of goods from the point of view of economic efficiency; (ii) the fragmentation of markets, as the economy produced too many small-scale goods, which resulted again in loss of efficiency; (iii) decreased competition from foreign firms and support for the monopoly power of domestic firms whose owners were politically well connected; and (iv) opportunities for rents and corruption, which raised input and transaction costs (Krueger 1974; Krugman 1993).

As government-led economic development strategies based on the structuralist teachings failed in many countries, the free market approach appeared to triumph and influence development thinking. This trend was reinforced by a new revolution in macroeconomics. The prevailing Keynesian macroeconomics was challenged by the stagflation in the 1970s, the Latin American debt crisis, and the collapse of the socialist planning system in the 1980s. The so-called rational expectations revolution emerged and refuted the structuralist theoretical foundation for the state’s role in using fiscal and monetary policy for economic development.

The Latin American debt crisis began in 1982 when international financial markets realized that the collapse of the Bretton Woods system had put some countries with unlimited access to foreign capital in a situation where they could not pay back their loans. The crisis was precipitated by a number of interrelated exogenous shocks that toppled the Mexican and several other Latin American economies, which were already overburdened with a substantial percentage of the world’s outstanding debt (Cardoso and Helwege 1995). It prompted multilateral lending institutions and bilateral lenders—especially the United States—to call for a comprehensive set of reforms of Latin American economies and to advocate a set of free-market policies that followed the canons of the neoclassical paradigm, later known as the Washington Consensus (Williamson 1990).

The Washington Consensus quickly came to be perceived as “a set of neoliberal policies that have been imposed on hapless countries by the Washington-based international financial institutions and have led them to crisis and misery” (Williamson 2002). It promoted economic liberalization, privatization, and the implementation of rigorous stabilization programs. The results of these policies in terms of growth and employment generation were at best controversial (Easterly, Loayza, and Montiel 1997; Easterly 2001). By the end of the 1990s and
parallel to the dismissal of structuralism and the prevalence of the free market approach, the development economics research community was witnessing the end of an era dominated by cross-country regressions, which attempted to identify growth determinants. That approach had been to focus on the independent and marginal effects of a multitude of growth determinants. This led to the linearization of complex theoretical models. Yet, the general view was that growth determinants interact with each other. To be successful, some policy reforms must be implemented with other reforms. There was a general perception that the policy prescriptions stemming from such regressions did not produce tangible results.

An alternative perspective on non-linearities was the Growth Diagnostics or Decision Tree approach suggested by Hausmann, Rodrik, and Velasco (2005). They recognized the central role of structural change in economic development and argued that there are “binding constraints” on growth in each country. These authors suggested that binding constraints can vary over time and across countries. They concluded that identification of the binding constraint was therefore key in practice. This framework highlighted pragmatically the inability of governments to reform everything and stressed the need to prioritize reforms, which should be done through the information revealed by shadow prices. It should be noted that the Growth Diagnostics approach is not operational unless one assumes away reform complementarities, which is the feature of linear growth regressions.

The divergence in growth performance between developed and developing countries, despite predictions of convergence from mainstream economic theory, has led to controversy. Some have concluded that the policy prescriptions, or expectations about their effectiveness, or both were wrong. Others have observed that growth researchers had paid limited attention to heterogeneity (the specific characteristics of each country). The suggestion that cross-country distribution may be multimodal (with the existence of “convergence clubs”) did not settle the debate about which new directions were needed for growth research. Instead, many basic questions have come back on the agenda: Are development economists looking in the wrong place in their quest for the determinants of growth? Should the focus be on institutions (institutional outcomes), instead of or in addition to policies? And, assuming that they are not reflecting other factors, how can good institutional outcomes be generated?

These unanswered questions were on the agenda for a long time. Starting in the 1980s, many development economists tried to understand better the causality of relationships and the various transmission channels through which policies, institutional changes, or foreign aid affect growth. They were also the rationale for an increased focus of growth research on microbehavior issues at the household and firm levels, with two goals: (i) allowing for heterogeneity in the
economy (across and within countries); and (ii) investigating how constraints to growth operate at the microlevel.

The growing disappointment and disillusionment with aid effectiveness also led to the quest for rigorous impact evaluation of development projects and programs. This has generated a new approach to development led by economists at the MIT Poverty Lab, whose goal is “to reduce poverty by ensuring that policy is based on scientific evidence” through the use of randomized control trials (RCT) or social experiments. Although RCT are good tools for understanding the effectiveness of some specific microprojects, they often do not start from a clear strategic assessment of how a particular method would fit the knowledge gaps of highest priority (Ravallion 2009). All too often, research looks for topics “under the light.” The positive outcomes for policymaking are more often the occasional by-products of research than its objective from the outset.

Recent microempirical studies may have indeed shed light on some important problems, such as the impact of the investment climate on firm performance or the impact of household behavior on productivity (Rosenzweig and Wolpin 1985). But “there is a risk the bulk of present-day research in development economics appears to be too narrowly focused and/or of too little generalizability to help much in the fight against poverty and to facilitate structural change and sustained growth” (World Bank 2010).

The time has come to reexamine the state of development economics, to learn from past experiences and previous knowledge, and to offer new thinking and a new framework. Drawing lessons from past experience and from economic theories, the next section presents the key principles of a new structural economics, which is a neoclassical approach to economic structure and dynamic change in the process of economic development.²

A Neoclassical Approach to Structure and Change

The starting point for the analysis of economic development is an economy’s endowments. Endowments are given in an economy at any specific time and are changeable over time. Following the tradition of classical economics, economists tend to think of a given country’s endowments as consisting only of its land (or natural resources), labor, and capital (both physical and human).³ These are in fact factor endowments, which firms in an economy can use in production. It should be noted that the analysis of new structural economics focuses on the dynamics of the capital/labor ratio. This is because land is exogenously given in any realistic discussion of a country’s development and natural resources, such as mining resources, exist underground in fixed quantity and their discovery is often random.
Conceptually, it is useful to add infrastructure as one more component in an economy’s endowments. Infrastructure includes hard (or tangible) infrastructure and soft (or intangible) infrastructure. Examples of hard infrastructure are highways, port facilities, airports, telecommunication systems, electricity grids, and other public utilities. Soft infrastructure consists of institutions, regulations, social capital, value systems, and other social, economic arrangements. Infrastructure affects the individual firm’s transaction costs and the marginal rate of return on investment.

Countries at different levels of development tend to have different economic structures due to differences in their endowments. Factor endowments for countries at the early levels of development are typically characterized by a relative scarcity of capital and relative abundance of labor or resources. Their production activities tend to be labor intensive or resource intensive (mostly in subsistence agriculture, animal husbandry, fishery, and the mining sector) and usually rely on conventional, mature technologies and produce “mature,” well-established products. Except for mining and plantations, their production has limited economies of scale. Their firm sizes are usually relatively small, with market transactions often informal, limited to local markets with familiar people. The hard and soft infrastructure required for facilitating that type of production and market transactions are limited and relatively simple and rudimentary.

At the other extreme of the development spectrum, high-income countries display a completely different endowment structure. The relatively abundant factor in their endowments is typically capital, not natural resources or labor. They tend to have comparative advantage in capital intensive industries with economies of scale in production. The various types of hard infrastructure (power, telecommunication, roads, port facilities, etc.) and soft infrastructure (regulatory and legal frameworks, cultural value systems, etc.) that are needed must comply with the necessities of national and global markets where business transactions are long distance and large in quantity and value.

Economic development requires continuous introduction of new and better technology to an existing industry. Most people in low-income countries depend on agriculture for their livelihood. Improvements in agricultural technology are key to increasing farmers’ income and reducing poverty. However, economic development also requires continuous diversifying and upgrading from existing industries to new, more capital-intensive ones. Without such a structural change, the scope for sustained increase in per capita income will be limited. Therefore, the discussion in this paper will focus mostly on issues related to industrial upgrading and diversification.

Developing countries have the advantage of backwardness in the upgrading process and a whole spectrum of industries with different levels of capital intensity available for them to choose. However, they must first upgrade their factor
endowment structure, which requires their stock of capital to grow more rapidly than the labor force (see Ju, Lin, and Wang 2009). When they move up the industrial ladder in the process of economic development, they also increase their scale of production—because of the indivisibility of capital equipment. Their firms become larger and need a bigger market, which in turn necessitates correspondent changes in power, transportation, financial arrangements, and other soft infrastructure.

The process of industrial upgrading and diversification also increases the level of risk faced by firms. As firms move closer to the global technology frontier, it becomes increasingly difficult for them to borrow mature technology from advanced countries. They increasingly need to invent new technologies and products and thus face more risk. The idiosyncratic risk of a firm has three components based on risk sources: technological innovation, product innovation, and managerial capacity. At the early level of development, firms tend to use mature technologies to produce mature products for mature markets. At that level, the main source of risk is the managerial ability of firms’ owner-operators. At a higher level of development, firms often invent new technologies to produce new products for new markets. In addition to managerial capacity, such firms face risks arising from the maturity of technology and markets. Therefore, while technological innovation, product innovation, and managerial capacity all contribute to the overall level of risk associated with firms, their relative importance varies greatly from one industry to another and from one level of economic development to another.

With changes in the size of firms, scope of the market, and nature of risk, along with the upgrading of the industrial structure, the requirements for infrastructure services, both hard and soft, also change. If the infrastructure is not improved simultaneously, the upgrading process in various industries may face the problem of x-inefficiency, a phenomenon discussed by Leibenstein (1957). Because the industrial structure in an economy at a specific time is endogenous to its given relative abundance of labor, capital, and natural resources at that time, the economy’s factor endowment will change with capital accumulation or population growth, pushing its industrial structure to deviate from the optimal determined by its previous level.4

When firms choose to enter industries and adopt technologies that are consistent with the comparative advantage determined by changes in the country’s factor endowments,5 the economy is most competitive.6 As competitive industries and firms grow, they claim larger domestic as well as international market shares and create the greatest possible economic surplus in the form of profits and salaries. Reinvested surpluses earn the highest return possible as well, because the industrial structure is optimal for that endowment structure. Over time, this approach allows the economy to accumulate physical and human capital,
upgrading the factor endowment structure as well as the industrial structure and making domestic firms more competitive over time in more capital and skill-intensive products.

Firms care about profits. For them spontaneously to enter industries and choose technologies consistent with the economy’s comparative advantage, the price system must reflect the relative scarcity of factors in the country’s endowment. This only happens in an economy with competitive markets (Lin 2009a; Lin and Chang 2009). Therefore, a competitive market should be the economy’s fundamental mechanism for resource allocation at each level of its development. That kind of \textit{comparative advantage-following} approach in economic development may appear to be slow and frustrating in countries with major poverty challenges. In reality, it is the fastest way to accumulate capital and upgrade the endowment structure, and the upgrading of industrial structure can be accelerated by better access to technology and industries already developed by and existing in more advanced countries. At each level in their development, firms in developing countries can acquire the technologies (and enter the industries) that are appropriate for their endowment structure, rather than having to reinvent the wheel (Gerschenkron 1962; Krugman 1979). This possibility to use off-the-shelf technology and to enter into existing industries is what has allowed some of the East Asian newly industrialized economies to sustain annual GDP growth rates of 8 and even 10 percent.

As a country climbs up the industrial and technological ladder, many other changes take place: the technology used by its firms becomes more sophisticated, and capital requirements increase, as well as the scale of production and the size of markets. Market transactions increasingly take place at arm’s length. A flexible and smooth industrial and technological upgrading process therefore requires simultaneous improvements in educational, financial, and legal institutions, and in hard infrastructure so that firms in the newly upgraded industries can reduce transaction costs and reach the production possibility frontier (Harrison and Rodríguez-Clare 2010). Clearly, individual firms cannot internalize all these changes cost effectively, and spontaneous coordination among many firms to meet these new challenges is often impossible. Change in infrastructure requires collective action or at least coordination between the provider of infrastructure services and industrial firms. For this reason, it falls to the government either to introduce such changes or to coordinate them proactively.

Successful industrial upgrading in responding to change in an economy’s endowment structure requires that the pioneer firms overcome issues of limited information regarding which new industries are the economy’s latent comparative advantages determined by the changing endowment structure. Valuable information externalities arise from the knowledge gained by pioneer firms in both success and failure. Therefore, in addition to playing a proactive role in the
improvements of soft and hard infrastructures, the government in a developing
country, like that in a developed country, needs to compensate for the information
externalities generated by pioneer firms (Rodrik 2004; Lin 2009a; Lin and Monga
2011; Harrison and Rodríguez-Clare 2010).7

What is “New” About the New Structural Economics?

Like all learning ventures, economic development thinking is bound to be a con-
tinuous process of amalgamation and discovery, continuity, and reinvention. The
existing stock of knowledge has been the result of many decades of work by think-
ers from various backgrounds and disciplines and has come to light through
several waves of theoretical and empirical research. It is therefore only natural
that the proposed new structural economics has some similarities and differences
with previous strands in the development economics literature. Its main value-
added should be assessed on the new policy insights it provides and the perti-
nence of the research agenda ahead.

Difference with Earlier Literature on Structural Change

Earlier thinking on structural change in the context of economic development is
mostly associated with Rostow (1990 [1960]) and Gerschenkron (1962). In
trying to understand how economic development occurs and what strategies can
be adopted to foster that process, the former suggested that countries can be
placed in one of five categories in terms of their level of growth: (i) traditional
societies, characterized by subsistence economy, with output not traded or even
recorded, the existence of barter, high levels of agriculture, and labor-intensive
agriculture; (ii) societies with preconditions to growth, where there is an increase
in capital use in agriculture, the development of mining industries, and some
growth in savings and investment; (iii) societies in take-off mode, with higher
levels of investment and industrialization, accumulation of savings, and a decline
in the share of the agricultural labor force; (iv) societies that drive to maturity
and where wealth generation enables further investment in value adding industry
and development—growth becomes self-sustaining, industry is diversified, and
more sophisticated technology is used; and (v) mass-consumption societies that
achieve high output levels and where the services industry dominates the
economy.

Gerschenkron questioned Rostow’s proposition that all developing countries
pass through a similar series of levels and its implication that it is possible to gen-
eralize the growth trajectory of different countries. For the new structural econ-
omics, economic development from a low level to a high level is a continuous
spectrum, not a mechanical series of five distinguished levels. Although the change in an economy’s industrial structure reflects the changes in that economy’s endowment structure, the development of industries in different countries with a similar endowment structure can be achieved in different and nonlinear ways. This is especially true with the increased globalization of markets, the rapid development of new products, and constant technological change, as countries can exploit opportunities that were not available in the past and specialize in industries that are likely to vary from one economy to another.

The new structural economics also provides a framework for understanding the endogeneity and exogeneity issues surrounding the key stylized facts of modern growth analysis that have been outlined by the Growth Commission (2008) and Jones and Romer (2009): an economy that follows its comparative advantage in the development of its industries will be most competitive in domestic and world markets. As a result, the economy will generate potentially the largest income and surplus for savings. Capital investment will also have the largest possible return. Consequently, households will have the highest savings propensity, resulting in an even faster upgrade of the country’s endowment structure (Lin and Monga 2010).

**Similarities and Differences with Old Structural Economics**

In terms of similarities, the “new” and the “old” structural economics are both founded on structural differences between developed and developing countries and acknowledge the active role of the state in facilitating the movement of the economy from a lower level of development to a higher one. However, there are profound differences between these two approaches regarding their targets and the modalities of state intervention. The old structural economics advocates development policies that go against an economy’s comparative advantage and advise governments in developing countries to develop advanced capital-intensive industries through direct administrative measures and price distortions. By contrast, the new structural economics stresses the central role of the market in resource allocation and advises the state to play a facilitating role to assist firms in the process of industrial upgrading by addressing externality and coordination issues.

The differences between the two frameworks derive from their dissimilar views on the sources of structural rigidities: old structural economics assumes that the market failures that make the development of advanced capital-intensive industries difficult in developing countries are exogenously determined by structural rigidities due to the existence of monopolies, labor’s perverse response to price signals, and/or the immobility of factors. By contrast, the new structural economics posits that the failure to develop advanced capital-intensive industries in developing countries is endogenously determined by their endowments. The
relative scarcity in their capital endowment and/or the low level of soft and hard infrastructure in developing countries make the reallocations from the existing industries to the advanced capital-intensive industries unprofitable for the firms in a competitive market.

Old structural economics assumes a dual and restrictive view of the world, with a binary classification of only two possible categories of countries: “low-income, periphery countries” versus “high-income, core countries.” As a result, it views the differences in the industrial structure between developed and developing countries as expressing a dichotomy. Contrary to that vision, the new structural economics considers these differences as the reflection of a whole spectrum that includes many different levels of development. The new structural economics also rejects dependency theories. In an increasingly globalized world, it sees opportunities for developing countries to counter negative historical trends by diversifying their economy and building industries that are consistent with their comparative advantage so as to accelerate growth and achieve convergence by exploiting the advantage of backwardness in an open, globalized world.

Another major difference between the new and the old structural economics is the rationale for using key instruments of economic management. Old structural economics sees systematic government intervention in economic activities as the essential ingredient in the modernization objective. Among the key instruments used to move from “developing” countries to “industrialized” countries are generalized protectionism (such as government-imposed tariffs on imports to protect infant industries), rigid exchange-rate policies, financial repression, and the creation of state-owned enterprises in most sectors.

By contrast, the new structural economics recognizes that import substitution is a natural phenomenon for a developing country climbing the industrial ladder in its development process, provided that it is consistent with the shift in comparative advantage that results from changes in its endowment structure. But it rejects conventional import-substitution strategies that rely on the use of fiscal policy or other distortions in low-income, labor or resource-abundant economies to develop high cost, advanced capital-intensive industries, which are not consistent with the country’s comparative advantage. It also stresses the idea that the industrial upgrading process in a developing country should be consistent with the change in the country’s comparative advantage that reflects the accumulation of human and physical capital and the change in its factor endowment structure—this ensures the viability of firms in new industries. The new structural economics concludes that the role of the state in industrial diversification and upgrading should be limited to the provision of information about the new industries, the coordination of related investments across different firms in the same industries, the compensation of information externalities for pioneer firms, and the nurturing of new industries through incubation and encouragement of
foreign direct investment (Lin 2009a; Lin and Chang 2009; Lin and Monga 2011). The state also needs to assume effectively its leadership role in the improvement of hard and soft infrastructure in order to reduce transaction costs on individual firms and so facilitate the economy’s industrial development process.

**New Structural Economics: Some Policy Insights**

The ultimate goal of development thinking is to provide policy advice that facilitates the quest for sustainable and inclusive economic and social progress in poor countries. Although specific policy measures to be derived from the new structural economics approach will require further research and depend very much on country context and circumstances, in this section I will make some conjectures about a few preliminary insights on various topics.

**Fiscal Policy.** Until Britain’s very high unemployment of the 1920s and the Great Depression, economists generally held that the appropriate stance for fiscal policy was for governments to maintain balanced budgets. The severity of the early twentieth-century crises gave rise to the Keynesian idea of counter-cyclicality, which suggested that governments should use tax and expenditure policies to offset business cycles in the economy. By contrast, neoclassical economics offers doubts about the implicit assumption behind the Keynesian model of a multiplier greater than one and its implication that governments are able to do something that the private sector has been unable to do: mobilize idle resources in the economy (unemployed labor and capital) at almost zero social cost, that is, with no corresponding decline in other parts of GDP (consumption, investment, and net exports). Instead, they warn against the possibility of the so-called Ricardian equivalence trap and point to the fact that households tend to adjust their behavior for consumption or saving on the basis of expectations about the future. They suggest that expansionary fiscal policy (stimulus packages) is perceived as immediate spending or tax cuts that will need to be repaid in the future. They conclude that the multiplier could be less than 1 in situations where the GDP is given and an increase in government spending does not lead to an equal rise in other parts of GDP. The neoclassical paradigm even suggests the possibility of some rare instances where multipliers are negative, pointing to situations where fiscal contractions become expansionary (Francesco and Pagano 1991).

From the viewpoint of new structural economics, the effects of fiscal policy may be different in developed and developing countries due to the differences in opportunities of using counter-cyclical expenditure for making productivity-enhanced investments. Physical infrastructure in general is a binding constraint for growth in developing countries and governments need to play a critical role in
providing essential infrastructure to facilitate economic development. In such contexts, recessions are typically good times for making infrastructure investments, for three main reasons. First, such investments boost short-term demand and promote long-term growth. Second, their investment cost is lower than in normal times. And third, the Ricardian equivalence trap can be avoided because the increase in future growth rates and fiscal revenues can compensate for the cost of these investments (Lin 2009b).

If a developing country government follows the new structural economics approach of facilitating the development of industries according to the country’s comparative advantage, its economy will be competitive and the fiscal position and the external account are likely to be sound, thanks to the likelihood of strong growth, good trade performance, and the lack of nonviable firms that the government has to subsidize. Under this scenario, the country will face fewer homegrown economic crises. If the economy is hit by external shocks such as the recent global crisis, the government will be in a good position to implement a counter-cyclical fiscal stimulus and invest in infrastructure and social projects. Such public investments can enhance the economy’s growth potential, reduce transaction costs on the private sector, increase the rate of return on private investment, and generate enough tax revenues in the future to liquidate the initial costs.

In addition to its different stance on fiscal stimulus, the new structural economics approach also offers a different strategy for managing natural resource wealth. In resource-abundant countries, it would recommend that an appropriate share of revenues from commodities be used to invest in human capital, infrastructure, social capital, and compensation for first movers in new nonresource sectors so as to facilitate the structural transformation. To accomplish this with the greatest effect, these resources should finance investment opportunities that remove binding constraints on industrial diversification and upgrading, especially in the infrastructure and education sectors. Microeconomic analyses show that even when factory floor costs are comparable, inefficiencies in infrastructure can make it impossible for poor countries to compete on international markets. Freight and insurance costs in African countries are 250 percent of the global average, with road freight delays two to three times as long as in Asia. Lacking financial resources and the appropriate policy frameworks, many of these countries are often unable to sustain much needed investment and maintenance expenditures. In such contexts, the effective fiscal strategy would not be to keep natural resource revenues in sovereign funds and invest in foreign equity markets or projects but, rather, to use a substantial portion of the revenues for financing domestic or regional projects that facilitate economic development and structural change—i.e. projects that stimulate the development of new manufacturing industries, diversify the economy, provide jobs, and offer the potential of continuous upgrading.11
Monetary Policy. Old structural economics suggested that monetary policy should be under government control (not independent central banks) and directed at influencing interest rates and even sector credit allocation. But it also acknowledged that many other factors that influence the investment demand-schedule in developing countries are too powerful for monetary policy alone to achieve sufficient levels of investment, channel resources in strategic sectors, and combat unemployment.

Building on lessons from the rational expectations revolution, neoclassical economists doubted the idea that monetary policy could be used to support industrial development. It recommended that its main goal be price stability, and advocated the use of short-term interest rates by independent central banks to maintain the general level of prices (or to control money supply growth), and not to stimulate economic activity and trigger inflation.

The new structural economics envisions the possibility of using interest rate policy in developing countries as a counter-cyclical tool and as an instrument to encourage infrastructure and industrial upgrading investments during recessions—measures that may contribute to productivity growth in the future. Monetary policy is often ineffective for stimulating investment and consumption in recessions and excess capacity situations in developed countries, especially when nominal interest rates hit the zero bound in a context of limited profitable investment opportunities, pessimistic expectations, high unemployment rates, low confidence about the future, and the likelihood of liquidity traps. It should be noted, however, that developing countries are less likely to encounter such liquidity traps. Even when faced with excess capacity in existing domestic industries, their scope for industrial upgrading and diversification is large. Their firms have incentives to undertake productivity-enhancing, industrial-upgrading investments during recessions if interest rates are sufficiently low. Furthermore, they tend to have many infrastructure bottlenecks. Lowering interest rates in such contexts would also encourage investments in infrastructure.

The objective of monetary policy should be much broader than traditionally conceived under neoclassical economics—in economic slumps, it should aim at encouraging investment that removes bottlenecks on growth. In practical terms, this implies not just that interest rates should be lowered in the slump, as would be the case in most circumstances under a standard Taylor rule. It also implies that monetary authorities should resort to temporary interest rate subsidies, flexible credit allocation rules, or similar time-bound devices, targeting infrastructure through development banks that are identified as binding constraints, preferably in specific geographic locations where the payoff is the largest and where political economy constraints can be more easily managed.
Financial Development. There is ample consensus that financial development is essential to sustaining economic growth. There is however much less agreement on the specific role it plays in that process. Starting with the observation that one of the major constraints facing developing countries was limited capital accumulation, old structural economics regarded the problems of the financial sector in underdeveloped economies as resulting from widespread market failures that could not be overcome by market forces alone.\textsuperscript{12} They recommended that governments adopt a hands-on approach in that process, mobilize savings, and allocate credit to support the development of advanced capital-intensive industries. This very often led to financial repression (McKinnon 1973; Shaw 1973). In some countries, especially in Sub-Saharan Africa, the belief in soft-budget constraints led governments to accumulate deficits in state-owned financial institutions and created a pervasive business culture of self-repression not only for banks, but also for private enterprises (Monga 1997). Drawing consequences from such analyses, neoclassical economists advocated financial liberalization. They contended that bureaucrats generally do not have the incentives or expertise to intervene effectively in credit allocation and pricing, and that a well-defined system of property rights, good contractual institutions, and competition would create the conditions for the emergence of a sound financial system. They recommended that government exit from bank ownership and lift restrictions on the allocation of credit and the determination of interest rates (Caprio and Honohan 2001).

While agreeing with the need to address the deleterious effects of financial repression, the new structural economics would emphasize the fact that those distortions are often designed to protect nonviable firms in priority sectors in developing countries. It would then stress the importance of an appropriate sequencing of liberalization policies in domestic finance and foreign trade so as to achieve stability and dynamic growth simultaneously during transition. The new structural economics also posits that the optimal financial structure at a given level of development may be determined by the prevailing industrial structure, the average size of firms, and the usual type of risk they face, all factors that are in turn endogenous to the economy’s factor endowments at that level. Observing that national policies frequently favor large banks and the equity market regardless of the structure of the economy, it would suggest that low-income countries choose small, local banks as the backbone of their financial systems, instead of trying to replicate the financial structure of advanced industrialized countries. This would allow small-scale firms in agriculture, industry, and the service sector to gain adequate financial services. As industrial upgrading takes place and the economy relies increasingly on more capital-intensive industries, the financial structure will change to give greater weight to large banks and sophisticated equity markets (Lin, Sun, and Jiang 2009).
Foreign Capital. In a world that they thought was characterized by the core–periphery relationship, old structural economists tended to view foreign capital mainly as a tool in the hands of industrialized countries and their multinational firms to maintain harmful control over developing countries. They rejected the idea that free capital movements among countries could deliver an efficient allocation of resources and considered foreign direct investment flows to poor countries as an instrument for foreign ownership and domination. They advocated tight restrictions on virtually all forms of international financial flows.

Neoclassical economic theory argues that international capital mobility serves several purposes: it allows countries with limited savings to attract financing for productive domestic investment projects; it enables investors to diversify their portfolios; it spreads investment risk more broadly; and it promotes intertemporal trade—the trading of goods today for goods in the future (Eichengreen and others 1999). Therefore, the theory generally favors open or liberalized capital markets, with the expectation of more efficient allocation of savings, increased possibilities for diversification of investment risk, faster growth, and the dampening of business cycles. It should be noted, however, that some neoclassical economists also argue that liberalized financial markets in developing countries can be distorted by incomplete information, large and volatile movements in and out of the system, and many other problems leading to suboptimal consequences that are damaging for general welfare.

The new structural economics approach considers foreign direct investment to be a more favorable source of foreign capital for developing countries than other capital flows because it is usually targeted toward industries consistent with a country’s comparative advantage. It is less prone to sudden reversals during panics than bank loans, debt financing, and portfolio investment, and does not generate the same acute problems of financial crises as do sharp reversals of debt and portfolio flows. In addition, direct investment generally brings technology, management, access to markets, and social networking, which are often lacking in developing countries and are yet crucial for industrial upgrading. Thus, liberalizing inward direct investment should generally be an attractive component of a broader development strategy. By contrast, portfolio investment that may move in and out quickly, in a large quantity, tends to target speculative activities (mostly in equity markets or the housing sector) and create bubbles and fluctuations. It should not be favored. The new structural economics approach may also shed new light on the puzzle raised by Lucas (1990) about the flow of capital from capital scarce developing countries to capital abundant developed countries. Without improvement of infrastructure and upgrading to new comparative advantage industries, the accumulation of capital in a developing country may encounter diminishing returns, causing lower returns to capital in developing countries, and justifying the subsequent outflow of capital to developed countries.
Trade Policy. There have been various old structural economics approaches to external trade. But one constant feature is the belief that integration into the global economy is bound to maintain the existing world power structure, with Western countries and their multinational corporations dominating poorer countries and exploiting their economies. In order to break the dependency trap, old structural economics thinkers have suggested that priority be given to import-substitution strategies, with developing economies closed and protected until their modern industries can compete with advanced industrialized countries in world markets.

A radically different view was adopted by economists in the 1980s. Observing that macroeconomic crises in developing countries almost always have an external dimension, they considered that their immediate cause was the lack of foreign exchange to service debts and purchase imports. They recommended trade liberalization and export promotion as a solution to generate foreign exchange through export earnings. This was also consistent with the view that, in the long term, outward oriented development strategies are more effective than inward looking policies. This view was bolstered further by the argument that such a strategy would increase demand for unskilled labor and hence unskilled wages, as had happened in successful East Asian countries (Kanbur 2009).

The analysis from the new structural economics would be consistent with the view from neoclassical economics that exports and imports are endogenous to the comparative advantage determined by a country’s endowment structure (they are essential features of the industrial upgrading process and reflect changes in comparative advantage). Globalization offers a way for developing countries to exploit the advantages of backwardness and achieve a faster rate of innovation and structural transformation than is possible for countries already on the global technology frontier. Openness is an essential channel for convergence. The new structural economics approach recognizes, however, that many developing countries start climbing the industrial ladder with the legacy of distortions from old structural economics strategies of import-substitution. It would therefore suggest a gradualist approach to trade liberalization. During transition, the state may consider providing some temporary protection to industries that are not consistent with a country’s comparative advantage, while liberalizing at the same time entry to other more competitive sectors that were controlled and repressed in the past. The dynamic growth in the newly liberalized sectors creates the conditions for reforming the old priority sectors. This pragmatic, dual-track approach may achieve the goal of growth without losers in the transition process (Naughton 1995; Lau, Qian, and Roland 2000; Subramanian and Roy 2003; Lin 2009a).
Human Development. Old structural economics generally said little about the role of human development in economic growth. By contrast, neoclassical economics has shown that the continuing growth in per capita incomes of many countries during the nineteenth and twentieth centuries was mainly due to the expansion of scientific and technical knowledge that raised the productivity of labor and other inputs in production. Economic theory has demonstrated that growth is the result of synergies between new knowledge and human capital, which is why large increases in education and training have accompanied major advances in technological knowledge in all countries that have achieved significant economic growth. Education, training, and health, which are the most important investments in human capital, are considered to be the most important driving force for economic development (Becker 1975; Jones and Romer 2009).

The new structural economics considers human capital to be one component of a country’s endowment. For economic agents, risks and uncertainty arise during the process of industrial upgrading and technological innovation that accompanies economic development. As various firms move up the industrial ladder to new, higher capital-intensity industries and get closer to the global industrial frontier, they face higher levels of risks. Human capital increases workers’ ability to cope with risk and uncertainty (Schultz 1961) but its formation requires a long time. A person who loses the opportunity to receive education at a young age may not be able to compensate for that loss at a later age. In a dynamic growing economy, it is important to plan ahead and make human capital investments before the economy requires the set of skills associated with new industries and technologies. However, improvements in human capital should be commensurable with the accumulation of physical capital and the upgrading of industry in the economy. Otherwise, human capital will either become a binding constraint for economic development if it is under-supplied because of insufficient investment, or the country will have many frustrated highly educated youths if the industrial upgrading of the economy is not progressing fast enough to provide skilled jobs.

A well-designed policy on human capital development should be an integral part of any country’s overall development strategy. The new structural economics goes beyond the neoclassical generic prescription for education and suggests that development strategies include measures to invest in human capital that facilitate the upgrading of industries and prepare the economy to make full use of its resources. The key components of such strategies should follow Lucas’s (2002) suggestion to allow human capital to have both a quality and a quantity dimension. It should also include alternative policies for promoting skill formation that are targeted to different levels of the life cycle, with the government and the private sector working closely together to anticipate or respond to the skills needs in the labor market. Singapore, one of the 13 high-growth economies that have
been able to grow at more than 7 percent for periods of more than 25 years since World War II, provides a successful example of human capital development as a national strategy (Osman-Gani 2004), which goes beyond the schooling decision and recognizes that on-the-job training is an important component of aggregate human capital. Its human resource strategies have been continuously revised and adjusted in conjunction with other national strategic economic policies.

Concluding Thoughts

The new structural economics approach highlights the importance of endowments and differences in industrial structures at various levels of development and the implications of distortions stemming from past, misguided, interventions by policymakers whose belief in old structural economics led them to overestimate governments’ ability to correct market failures. It also points out the fact that policies advocated under the Washington Consensus often failed to take into consideration the structural differences between developed and developing countries and ignored the second-best nature of reforming various types of distortions in developing countries.

The proposed new structural economics attempts to develop a general framework for understanding the causality behind the observed stylized facts of sustained growth. Specifically, the new structural economics proposes to: (i) develop an analytical framework that takes into account factor and infrastructure endowments, the levels of development, and the corresponding industrial, social, and economic structures of developing countries; (ii) analyze the roles of the state and the market at each development level and the mechanics of the transition from one level to another; and (iii) focus on the causes of economic distortions and the government’s strategies for exit from the distortions. It is not an attempt to substitute another ideologically based policy framework for those that have dominated development thinking in past decades, yet showing little connection to the empirical realities of individual countries. Rather, it is an approach that brings attention to the endowment structure and level of development of each country and suggests a path toward country-based research that is rigorous, innovative, and relevant to development policy. This framework stresses the need to understand better the implications of structural differences at various levels of a country’s development—especially in terms of the appropriate institutions and policies, and the constraints and incentives for the private sector in the process of structural change.

The current state of development economics and the severe impact of the global crisis on the economies of developing countries have generated strong demand for a new framework for development thinking. The research agenda of
the new structural economics should enrich research and enhance the understanding of the nature of economic development. This would help assist low and middle-income countries in achieving dynamic, sustainable, and inclusive growth, and in eliminating poverty.

Acknowledgements


Notes

Justin Yifu Lin is Senior Vice President and Chief Economist of the World Bank; email address: justinlin@worldbank.org.

1. The paper was presented as the Kuznets Lecture at the Economic Growth Center, Yale University on March 1, 2011. The main arguments of this paper were first presented at DEC’s fourth Lead Economists Meeting and at Lin’s first anniversary at the Bank on June 2, 2009. A shorter version of the paper was presented at the conference on “Challenges and Strategies for Promoting Economic Growth,” organized by the Banco de México in Mexico City on October 19–20, 2009, and at public lectures in Cairo University on November 5, 2009, Korean Development Institute on November 17, 2009, OECD on December 8, 2009, UNU-WIDER on January 19, 2010, Stockholm Institute of Transitional Economics on January 21, 2010, National University of Management in Cambodia on September 8, 2010, Bank of Italy on April 26, 2011, and University of Dar es Salaam on April 29, 2011.

2. I will refer to the early contributions by structuralist economists such as Prebisch (1950) and Furtado (1964, 1970) and recent contributions by structuralist economists such as Taylor (1983, 1991, 2004) and Justman and Gurion (1991) as old structural economics.

3. The total endowments at a specific time—the economy’s total budgets at that time and the endowment structure, together with the households’ preferences and firms’ available production technologies—determine the relative factor and product prices in the economy. Total budgets and relative prices are two of the most fundamental parameters in economic analysis. Moreover, the endowments are given at any specific time and are changeable over time. These properties make endowments and the endowment structure the best starting point for analysis of economic
development. Except in Heckscher-Ohlin trade theory, the economic profession has not given sufficient attention to the implications of factor endowments and endowment structure.

4. The proposition that the industrial structure is endogenous to an economy’s endowment structure at each level of its development has been the subject of extensive theoretical studies. For instance, Lin and Zhang (2009) develop an endogenous growth model that combines structural change with repeated product improvements to discuss the endogeneity of industrial structure, the appropriate technology, and economic growth in a less developed country (LDC) in a dynamic general-equilibrium framework. They use a two-sector model in which technological change in the traditional sector takes the form of horizontal innovation based on expanding variety as suggested in Romer (1990) while technological progress in the modern sector is accompanied by incessantly creating advanced capital-intensive industry to replace backward labor-intensive industry. This requires an intentional investment of resources by profit-seeking firms or entrepreneurs (Grossman and Helpman 1994). The model shows that: (i) the optimal industrial structure in LDCs should not be the same as that in developed countries (DCs); (ii) the appropriate technology adopted in the modern sector in LDCs ought to be inside the technology frontier of the DCs; and (iii) a firm in an LDC that enters capital-intensive, advanced industry (by DC standards) would be nonviable owing to the relative scarcity of capital in the LDC’s factor endowment. Ju, Lin, and Wang (2009) develop a dynamic general equilibrium model to show that industries will endogenously upgrade toward the more capital-intensive ones as the capital endowment becomes more abundant. The model features a continuous inverse-V-shaped pattern of industrial evolution driven by capital accumulation: As the capital endowment reaches a certain threshold, a new industry appears, prospers, then declines, and finally disappears. While the industry is declining, a more capital-intensive industry appears and booms.

Capital is mobile in an open economy. It is unlikely that the mobility of capital will equalize the capital–labor ratio in high-income, capital-abundant countries and low-income, labor-abundant countries. This is because there are two main purposes for the capital to flow from a higher-income country to a lower-income country. The first one is to exploit the lower-income country’s comparative advantage of abundant labor (or natural resources) so as to use the lower-income country as its export base. For this purpose, the industry must be consistent with the recipient, lower-income country’s comparative advantage determined by its factor endowment, although the technology used by the foreign-invested firms may be somewhat more capital intensive than the indigenous firms. The second purpose of capital flow from a higher-income country is to get access to a lower-income country’s domestic markets. For this type of capital flow, the foreign-invested industries will be more capital intensive than the indigenous firms but only the types of production activities that are consistent with the host country’s comparative advantage, for example assembly of parts into final products, will be located in the lower-income country. Therefore, the theoretical insights derived from the assumption that the relative abundance of capital in a country is given at any specific time will hold even with capital mobility.

5. For nontradable goods and services, the nature of least-cost production technology will also be endogenously determined by the endowment structure. That is, as capital becomes relatively abundant, the technology used to produce nontradable goods and services will also become relatively capital intensive, just as what happens in the tradable goods sector. For simplicity, the discussion in the paper will focus on the tradable sector.

6. Porter (1990) made the term “competitive advantage” popular. According to him, a nation will have competitive advantage in the global economy if the industries in the nation fulfill the following four conditions: (1) their industries intensively use the nation’s abundant and relatively inexpensive factors of production; (2) their products have large domestic markets; (3) each industry forms a cluster; and (4) the domestic market for each industry is competitive. The first condition in effect means that the industries should be the economy’s comparative advantage determined by the nation’s endowments. The third and the fourth conditions will hold only if the industries are consistent with the nation’s competitive advantage. Therefore, the four conditions can be reduced to two independent conditions: comparative advantage and domestic market size. Of these two independent conditions, comparative advantage is the most important because if an industry
corresponds to the country’s comparative advantage, the industry’s product will have a global market. That is why many of the richest countries of the world are very small (Lin and Ren 2007).

7. Industries in advanced developed countries today are typically located on the global frontier and face uncertainty as to what the next frontier industries will be. This explains why government policy measures to support pioneer firms in such countries are usually in the form of general support to research in universities (which has externalities to private firms’ R&D), patents, preferential taxes for capital investments, mandates, defense contracts, and government procurement. Support in the form of preferential taxes, defense contracts, and government procurement are industry or product-specific. Government support to basic research also needs to be prioritized for certain types of potential industries or products because of budget constraints. However, government support to pioneer firms in developing countries, especially low-income countries, often fails. One of the most important reasons is the attempt by low-income countries governments to support firms in industries that are inconsistent with the economy’s comparative advantages (Lin 2009a; Lin and Chang 2009).

8. Barro (2009) calls active fiscal policy of the Keynesian type “the extreme demand-side view” or the “new voodoo economics.”

9. Recent research suggests that economic returns on investment projects in developing countries averages 30–40 percent for telecommunications, more than 40 percent for electricity generation, and more than 200 percent for roads. In Thailand, production loss due to power outages represented more than 50 percent of the total indirect costs of doing business in 2006. Firms often rely on their own generators to supplement the unreliable public electricity supply. In Pakistan, more than 60 percent of firms surveyed in 2002 owned a generator. The cost of maintaining a power generator is often high and burdensome, especially for small and medium-size firms, which are important sources of employment. Yet, while these costs must be privately borne, their benefits are felt across the economy.

10. This is percentage of cost (UNCTAD Statistical Database).

11. The exploitation of natural resources can generate a large amount of revenues but it is generally very capital intensive and provides limited job opportunities. In a recent visit to Papua New Guinea, I observed that the Ok Tedi copper and gold mine in Tabubil generates almost 80 percent of the country’s export revenues and 40 percent of government revenues but provides only 2,000 jobs. A proposed liquefied natural gas project will double Papua New Guinea’s national income after its completion in 2012, but the project will only provide 8,000 jobs. The majority of Papua New Guinea’s 6.5 million population still live on subsistence agriculture. The contrast between the standard of living of a few elite workers in modern mining and that of subsistence farmers is becoming a source of social tensions. A similar observation can be made about Botswana: the failure to diversify the economy from diamond mining and to generate employment opportunities may explain the widening disparity and deterioration of various human and social indicators, despite the diamond industry’s great success in sustaining Botswana’s growth miracle over the past 40 years.

12. Gerschenkron (1962) made a similar point, arguing that the private sector alone cannot effectively address the problems of access to finance in weak institutional environments.

13. A sudden large inflow of portfolio capital is most likely to be invested in speculative sectors rather than in productive sectors. The reason is twofold: a large increase in investment in existing industries may encounter diminishing returns to capital, and the potential for quick and large industrial upgrading is limited by human capital, as well as soft and hard infrastructure constraints.

14. Carneiro and Heckman (2003) have demonstrated the importance of both cognitive and noncognitive skills that are formed early in life in accounting for gaps in schooling among social groups and other dimensions of socioeconomic success. They have provided empirical evidence of a high return to early interventions and a low return to remedial or compensatory interventions later in life.

15. The list includes: Botswana; Brazil; China; Hong Kong SAR, China; Indonesia; Japan; Korea; Malaysia; Malta; Oman; Singapore; Taiwan, China; and Thailand.
References

The word processed describes informally reproduced works that may not be commonly available through libraries.


Comments on “New Structural Economics” by Justin Yifu Lin

Anne Krueger

Ever since development economics became a field, there has been a search for “the” key to development. Physical capital accumulation, human capital, industrial development, institutional quality, social capital, and a variety of other factors have been the focus at one time or another. As each became the focal point, there was a parallel explicit or implied role of government.

If I understand Justin Lin correctly, he is saying that the “new structural economics” (NSE) accepts that earlier thought ignored comparative advantage, which should be market determined, but that growth requires improvements in ‘hard’ (tangible) and ‘soft’ (intangible) infrastructure at each stage. Such upgrading and improvements require coordination and inhere with large externalities to firms’ transaction costs and returns to capital investment. Thus, in addition to an effective market mechanism, the government should play an active role in facilitating structural change (p. 206).

He seems also to believe that growth depends almost entirely on industry growth and believes that constant “upgrading” or moving up the value added chain is the central challenge. He says that “the laissez-faire approach ... missed the importance of the process of continuous, fundamental technological changes and industrial upgrading, which distinguishes modern economic growth from premodern economic growth” (p. 196).

It is questionable whether such changes and upgrading must take place early in the development process. In many countries, unskilled labor has moved to unskilled-labor-intensive industries, with expansion of those industries’ outputs for a period during which more and more workers acquired acquaintance with modern factory techniques, and exports of the unskilled-labor-intensive goods...
increased. Only later in the development process did upgrading become a major part of industrial growth once there had been significant absorption of rural labor, and much of it happened in existing firms in response to rising real wages, lower capital costs, and learning through exposure to the international market.

However, in most countries rural labor could be absorbed only as agricultural productivity increased; Lin’s NSE seems to equate growth with industrial expansion, ignoring the importance of increased productivity of the large fraction of the labor force (and of land) in rural areas. Failure to invest in agricultural research and development and in rural health and education has been a major weakness of many countries’ development strategies. While strides have been made in reducing discrimination against agriculture, the NSE as expositied by Lin would appear to support the industrial and urban bias that has itself constituted a very large distortion in some countries.

It will come as no surprise that I agree that the market should be used to determine comparative advantage, and that governments have responsibilities for insuring an appropriate incentive framework and provision of infrastructure (both hard and, as he terms it, “soft”).

But there is nothing new in that. What purports to be the “new” part is the assertion that coordination and upgrading of infrastructure should in some way be related to particular industries. It is at this point where a question arises: most economists would accept the view that cost–benefit analysis should be used in the choice of infrastructure projects. If “externalities” and “coordination” are important, are they important for specific industries or for the entire industrial economy? If the former, how are those industries to be identified, and how would the externalities be estimated in cost–benefit analysis? Or would they? If infrastructure is seen to be industry-specific, it is not clear what it is. As with the possible existence of infant industries, it is one thing to believe that there are such industries (perhaps) and quite another to identify ahead of time which they are. And even if such industries exist and are identified, questions arise as to the incentives that would be appropriate for the government to foster these industries. (Would they be firm-specific treatment? Tariffs? Subsidies to firms or industries? Each has huge problems.) And if it is more “conventional,” what is new? If infrastructure is specific to industry (or a group of industries), the same questions must be addressed.

Some hints are given as to what Lin has in mind: “successful industrial upgrading in responding to change in an economy’s endowment structure requires that the pioneer firms overcome issues of limited information regarding which new industries are the economy’s latent comparative advantages determined by the changing endowment structure. Valuable information externalities arise from the knowledge gained by pioneer firms in both success and failure. Therefore, in addition to playing a proactive role in the improvements of soft and hard...
infrastructures, the government in a developing country, like that in a developed country, needs to compensate for the information externalities generated by pioneer firms” (p. 203).

Here, the infant industry concerns arise again. How can these externalities be forecast? As Baldwin (1969) pointed out, there are major difficulties with this argument, quite aside from the identification of such externalities. And firms producing unskilled-labor-intensive goods and exporting them have usually learned of the opportunities provided by the international market and chosen to upgrade as their experience has increased. Learning does not seem to have been a major issue for firms in South Korea, Taiwan, and elsewhere.

Another hint as to what Lin has in mind comes from his advocacy of coordination of infrastructure investments. According to him, “Change in infrastructure requires collective action or at least coordination between the provider of infrastructure services and industrial firms. For this reason, it falls to the government neither to introduce such changes or to coordinate them proactively.” (p. 203) How this would be carried out is unclear; Lin insists that infrastructure must be upgraded with growth as long as it is consistent with the evolving future direction of comparative advantage, but does not elaborate on how that future direction should be identified. Involving individual firms and industries in decisions as to infrastructure investments would appear to offer far too much scope for individual firms’ and industries’ influence over these investments.

Although it is certainly true that not everything can be done at once, focus on selected areas for large investments at the neglect of the rest of the economy is a highly questionable strategy. Why it would be preferable to allocate scarce capital so that some activities have excellent infrastructure while others must manage with seriously deficient infrastructure is not clear: without further evidence, it would appear to be a distortion. Further, questions can also be raised as to why “soft infrastructure,” such as the “business environment” (which consists of such things as the commercial code, the structure of taxes and subsidies, regulations, and so on), cannot be economy wide. And the criteria by which there would be designation of a given area, or the types of industries that would be eligible, as the recipient of special treatment are not discussed. What the hard infrastructure is that does not consist of items such as roads and ports, and is industry specific, is not discussed.

But all of this hinges on the proposition that decisionmakers in the public sector can ascertain the appropriate rate of “upgrading” and the extent of the supposed externalities. This raises a host of issues. There is, first, the consideration that even if one could know which activities would have comparative advantage, that advantage often develops as small firms enter, some of which are successful and grow larger. Any strategy of “upgrading” would inevitably favor larger, established firms, and hence encounter the same sorts of problems as did the older
import-substitution strategy which, as Lin recognizes, failed. “Picking winners” as industries is difficult; it cannot be firm specific or the usual problems of corruption and cronymism arise. And yet supporting an industry or industries as an undifferentiated entity is difficult: are textiles an industry? Or is synthetic fiber an industry? Or is nylon an industry? And, of course, the breakdown could go further. And as capital and skills per person accumulate, how is it to be decided where the industrial park or export processing zone should be? And which firms should be eligible to enter it?

Another strand of Lin’s argument pertains to the role of distortions. He appears to be saying that countries that earlier adopted import-substitution strategies have distorted industrial structures that should affect policy. In particular, he says: “many developing countries start climbing the industrial ladder with the legacy of distortions from old structural strategies of import-substitution. [The new structural economics] would therefore suggest a gradualist approach to trade liberalization. During transition, the state may consider some temporary protection to industries that are not consistent with a country’s comparative advantage, while liberalizing at the same time entry to other more competitive sectors that were controlled and repressed in the past” (p. 212).

Here, as elsewhere, little guidance is given as to how much protection industries would be provided with; how long that protection would last; how industries to be protected would be chosen; and so on. But even more important, one can imagine the political pressures for greater protection for longer periods. Protection of some industries is dis-protection of others, as is well known, so reform efforts would clearly be dampened. Even worse, a major challenge for liberalizing reforms is for it to be credible that the altered policies are not reversible. Lin’s prescription would greatly increase the challenge of creating credibility, and a slower transition would be a longer period during which growth was slow and political pressures opposing liberalization at all were mounting.

In all, there is much in Lin’s analysis with which most would agree, but focus on governmentally led identification of industries with “latent comparative advantage” and industry-specific provision of infrastructure is not convincing. Lin calls for much research. A first task should be to show that there are industry (or industry-cluster) externalities, how they could be identified and measured ex ante, and what sorts of government support would improve potential welfare and growth prospects without generating the same sorts of rent-seeking opportunities as import substitution policies did.

Until that research is undertaken, the NSE will, it is to be feared, be taken as a license for governments to support specific industries (and worse yet, perhaps even firms), in ways that may be no more conducive to growth than were the old, failed, import-substitution policies.
Anne Krueger is Professor of International Economics in the School of Advanced International Studies (SAIS) at Johns Hopkins University and a Senior Fellow at the Stanford Center for International Development.

Reference

Justin Lin wants to make structuralist economics respectable again, and I applaud him for that. He wants to marry structuralism with neoclassical economic reasoning, and I applaud this idea too. So he has two cheers from me. I withhold my third cheer so I can quibble with some of what he writes.

The central insight of structuralism is that developing countries are qualitatively different from developed ones. They are not just radially shrunk versions of rich countries. In order to understand the challenges of under-development, you have to understand how the structure of employment and production—in particular the large gaps between the social marginal products of labor in traditional versus modern activities—is determined and how the obstacles that block structural transformation can be overcome.

The central insight of neoclassical economics is that people respond to incentives. We need to understand the incentives of, say, teachers to show up for work and impart valuable skills to their students or of entrepreneurs to invest in new economic activities if we are going to have useful things to say to governments about what they ought to do. (And of course, let’s not forget that government officials must have the incentive to do the economically “correct” things too.)

If we put these two sets of ideas together, we can have a useful development economics, one that does not dismiss the tools of contemporary economic analysis and yet is sensitive to the specific circumstances of developing economies. This is the kind of development economics that is appropriately nuanced in its take on government intervention. It doesn’t presume omniscience or altruism on the part of governments. It has a healthy respect for the power and effectiveness of markets. But it does not blithely assume that development is an automatic process that takes care of itself as long as government stays out of the picture.
So as Lin rightly emphasizes, the state has a useful role to play in promoting industrial diversification and upgrading. He lists among desirable functions the provision of information about new industries, the coordination of investments across firms and industries, the internalization of informational externalities, and the incubation of new industries through the encouragement of foreign direct investment. Policies of this kind may be unnecessary or superfluous in advanced economies, but they are essential if poor countries are to progress.

To distinguish his brand of structuralist development economics from old-style structuralism Lin writes that a key difference is that the old school advocated policies that go against an economy’s comparative advantage. The new approach, by contrast, “stresses the central role of the market . . . and advises the state to play a facilitating role to assist firms in the process of industrial upgrading by addressing externality and coordination issues.” Lin argues that government policies should “follow” comparative advantage, rather than “defy” it.

Here is where I quibble with Lin’s argument. It seems to me that Lin wants to argue both for and against comparative advantage at the same time, and I cannot quite see how this can be done. If one believes that externality and coordination problems need to be addressed, as Lin apparently does, one must believe that such problems are preventing firms from investing appropriately. One must believe that markets are sending entrepreneurs the wrong signals—invest here, not there—and that allocating resources according to comparative advantage, as revealed by market prices, would be socially suboptimal. Comparative advantage has practical meaning for firms only insofar as it gets reflected in prices.

So when Lin asks governments to step in to address market failures and recommends the type of policies I have listed above—the coordination of investments, the incubation of new industries, etc.—he too is asking them to defy comparative advantage as revealed in market prices. In this respect, there is less difference between what the old school said and what the new school is saying.

Lin doesn’t want governments to employ “conventional” import substitution strategies to build capital-intensive industries which “are not consistent with the country’s comparative advantage.” But isn’t building industries that defy comparative advantage what Japan and South Korea did, in their time? Isn’t it what China has been doing, and quite successfully, for some time now? According to my calculations, the export bundle of China is that of a country between three and six times richer. If China, with its huge surplus of agricultural labor, were to specialize in the type of products that its factor endowments recommend, would it now be exporting the advanced products that it is?

Some people draw a distinction between static and dynamic comparative advantage in this context, but I don’t think that is the relevant distinction. Market failures drive a wedge between market prices and social marginal valuations, and distort the relative costs that signal comparative advantage. Whether these distortions are introduced into intertemporal relative prices or today’s
relative prices is largely secondary. The policies that Lin recommends are meant to offset such market distortions, and their intended effect is to induce firms to make choices that defy comparative advantage.

I suspect that my difference with Lin is mainly methodological—and perhaps even just terminological—and may have little practical import. What Lin probably has in mind is that today’s industrial policies need to have a softer touch than that which structuralists of old tended to recommend. They must be more respectful of markets and incentives; they must show greater awareness of the potential of government failures; and they must focus specifically on market failures rather than vague shortcomings of the private sector. I would agree with all this.

But a deeper question relates to the policy implications one draws from all this. In principle, market failures need to be addressed with appropriately targeted policies. So if the problem is one of information spillovers, the first-best is to subsidize the information generating process. If the problem is lack of coordination, the first-best is for the government to bring the parties together and coordinate their investments. In practice, though, the relevant market failures cannot be always closely identified and the directly targeted remedies may not be available. The practical reality is that the type of policies structuralism calls for—whether of the traditional or the contemporary type—have to be applied in a second-best setting. And in such a setting, nothing is all that straightforward anymore.

Presumably this is the reason why Lin recommends, for example, a gradual approach to trade liberalization. Such an approach is, at best, a second-best remedy to some loosely specified market failures that either cannot be precisely identified ex ante or cannot be fully treated with first-best Pigovian interventions. But how different is this from the old structuralist approach? Didn’t most structuralists also view protection as a temporary expedient, to be done away with once the requisite industrial capabilities were built?

To repeat, my differences with Justin Lin are second order, and they are swamped by our areas of agreement. My quibbles are a bit like the internal doctrinal debates waged among communists—does the revolution require the intensification of the class struggle, or can that stage be skipped?—when much of the rest of the world is on a different wavelength altogether.

As a fellow traveler, I am greatly encouraged by what Justin Lin is trying to do. It is high time that the common sense exhibited in his approach reclaimed its mantle in development economics.

Note

Dani Rodrik is professor of international political economy at Harvard University’s John F. Kennedy School of Government (dani_rodrik@harvard.edu).
Twelve years ago, when I was chief economist of the World Bank, I suggested that the major challenge to development economics was learning the lessons of the previous several decades: a small group of countries, mostly in Asia, but a few in other regions, had had phenomenal success, beyond anything that had been anticipated by economists; while many other countries had experienced slow growth, or even worse, stagnation and decline—consistent with the standard models in economics which predicted convergence. The successful countries had followed policies that were markedly different from those of the Washington Consensus, though they shared some elements in common; those policies had not brought high growth, stability, or poverty reduction. Shortly after I left the World Bank, the crisis in Argentina—which had been held up as the poster child of the country that had followed Washington Consensus policies—reinforced the doubts about that strategy.

The global financial crisis, too, has cast doubt over the neoclassical paradigm in advanced industrial countries, and rightly so. Much of development economics had been viewed as asking how developing countries could successfully transition toward the kinds of market-oriented policy frameworks that came to be called “American style capitalism.” The debate was not about the goal, but the path to that goal, with some advocating “shock therapy,” while others focused on pacing and sequencing—a more gradualist tack. The global financial crisis has now raised questions about that model even for developed countries.

In this short essay, I want to argue that the long-term experiences in growth and stability of both developed and less developed countries, as well as the deeper theoretical understanding of the strengths and limitations of market economies, provide support for a “new structural” approach to development—an approach...
similar in some ways to that advocated by Justin Lin in his paper, but markedly different in others. This approach sees the limitations of markets as being greater than he suggests—even well functioning market economies are, on their own, neither efficient nor stable. The only period in the history of modern capitalism when there has not been repeated financial crises was the short period after the Great Depression when the major countries around the world adopted, and enforced, strong financial regulations. Interestingly this was also a period of rapid growth and a period in which the fruits of that growth were widely shared.

But government not only has a restraining role; it has a constructive and catalytic role—in promoting entrepreneurship, providing the social and physical infrastructure, ensuring access to education and finance, and supporting technology and innovation.

The perspective that I am putting forward differs not only in its view of the efficiency and stability of unfettered markets, but also in what it sees as the primary driver of economic growth. Since Solow’s pioneering work more than a half-century ago (Solow 1957), it has been recognized that the major source of increases in per capita income are advances in technology.¹

The argument that improvements in knowledge are a primary source of growth is even more compelling for developing countries. As the World Development Report for 1998–99 emphasized, what separates developing and developed countries is not just a gap in resources, but a disparity in knowledge. There are well understood limits to the pace with which countries can accumulate capital, but the limitations on the speed with which the gap in knowledge can be closed are less clear.

But the view that creating a learning society, focusing on absorbing and adapting, and eventually producing knowledge, provides markedly different perspectives on development strategies than those provided by the neoclassical model. That model centered attention on increasing capital and the efficient allocation of resources. Since the appropriate sectoral structure of the economy naturally depends on the resource endowment, there will be a natural evolution of the economy’s structure over time. Markets allocate resources efficiently, enabling the structure to change as the (endogenous) endowments change. A government’s main role, in this view, is not to put impediments in the market.

The standard market failures approach criticized these conclusions by focusing on a variety of market imperfections: For instance, imperfections in capital markets meant that finance was often not available for new enterprises that were required as part of this sectoral adjustment. Individuals on their own couldn’t finance their education. There are pervasive externalities—not only environmental externalities but also those associated with systemic risk, so evidenced in the current crisis. Research over the past 20 years has explored the consequences of market failures like imperfect capital markets, traced these imperfections back
to problems of imperfect and asymmetric information, and proposed a set of reme-
dies, which in some countries, in some periods, have worked remarkably well. Good financial regulations in countries like India protected them against the ravages of the global financial crisis.

But the perspective of the “learning society”—or, as Greenwald and I call it, the “infant economy”—adds a new dimension to the analysis (Greenwald and Stiglitz 2006). Knowledge is different from an ordinary commodity. The accumulation of knowledge is inherently associated with externalities—knowledge spillovers. Knowledge itself is a public good. If the accumulation, absorption, adaptation, production, and transfer of knowledge are at the center of successful development, then there is no presumption that markets, on their own, will lead to successful outcomes. Indeed there is a presumption that they will not.

The “new structuralist approach” advocated by Justin Lin is perfectly aligned with this perspective. Lin provides guidance as to how governments should direct the economy; he emphasizes that they should strive to shape the economy in a way that is consistent with its comparative advantage. The problem is that some of the most important elements of comparative advantage are endogenous. Switzerland’s comparative advantage in watch-making has little to do with its geography.

Standard Heckscher-Ohlin theory (emphasizing that trade in goods was a substitute for movement in factors) was formulated in a period before globalization allowed the kinds of flows of capital that occur today. With fully mobile capital, outside of agriculture, natural resource endowments need not provide the basis for explaining patterns of production and specialization. In short, there is no reason for countries to need to limit themselves to patterns dictated by endowments, as conventionally defined. More important is the “endowment” of knowledge and entrepreneurship. A major focus of policy should be on how to enhance and shape those endowments.

Even if a government would like to avoid addressing these issues, it cannot; for what the government does (or does not do) has consequences, positive and negative, for the development of the “learning society.” This is obviously so for investments in infrastructure, technology, and education; but also for financial, trade, intellectual property rights and competition policies.

At the center of creating a learning society is the identifying of sectors that are more amenable to learning, with benefits not captured by firms themselves, so that there will be underinvestment in learning. Elsewhere Greenwald and I have argued that an implication of this is the encouragement of the industrial sector, which typically has large spillovers. This approach provides an interpretation of the success of Asia’s export-led growth. Had Korea allowed market forces on their own to prevail, it would not have embarked on its amazing development successes. Static efficiency entailed that Korea produced rice; indeed the country
might today have been among the most efficient rice farmers—but it would still be a poor country. As Arrow pointed out (1962), one learns by doing (and one learns how to learn by learning [Stiglitz 1987]).

This discussion highlights the fundamental difference with neoclassical approaches emphasizing short-run efficiency. The fundamental trade-offs between static and dynamic efficiency should be familiar from the debate over patent laws.

A major concern with these industrial policies concerns implementation—do developing countries have the requisite capacities? We need to put this question in context.

There is probably no country that has grown successfully without an important role, not just in restraining and creating markets, but also in promoting such industrial policies, from the countries of East Asia today to the advanced industrial countries, not just during their developmental stages, but even today. The task is to adopt policies and practices—to create institutions like an effective civil service—that enhance the quality of the public sector. The successful countries did so. Policies that either intentionally or unintentionally weaken the state are not likely to do so.

Economic policies have to reflect the capacity of the state to implement them. One of the arguments in favor of exchange rate policies that encourage export industries is that they are broad based: the government does not have to pick particular “strategic” sectors to support. As always, there are trade-offs: efficiency might be enhanced if the sectors with the largest externalities could be targeted.

There are other broad-based policies, such as a development-oriented intellectual-property regime, and investment and financial policies that encourage transfer of technology and the promotion of local entrepreneurship, that can help promote a learning and innovation society (Hausman and Rodrik 2003; Stiglitz 2004; Emran and Stiglitz 2009; Hoff 2010). Some forms of financial and capital market liberalization may be counterproductive.

Interventions will never be perfect, nor need they be to effect an improvement in economic performance. The choice is not between an imperfect government and a perfect market. It is between imperfect governments and imperfect markets, each of which has to serve as a check on the other; they need to be seen as complementary, and we need to seek a balance between the two—a balance which is not just a matter of assigning certain tasks to one, and others to the other, but rather designing systems where they interact effectively.

While I have been discussing the economics of development, that subject cannot be separated from broader aspects of societal transformation (Stiglitz 1998), as Hirschman emphasized in his writings (1958, 1982). Race and caste are social constructs that effectively inhibit the human development of large parts of the population in many parts of the world. The study of how these constructs get formed, and how they change, is thus a central part of developmental studies.
(Hoff and Stiglitz 2010). In this article, I have emphasized the creation of a learning society. The economics of doing so entails policies that change sectoral composition. But at the root of success is the education system and how it inculcates attitudes toward change and skills of learning. Other policies (for example legal systems, gender-based microcredit schemes, affirmative action programs) can also play an important role.

Before concluding, I want to make two further remarks. The first concerns the relationship between growth and poverty reduction. While growth may be necessary for sustained poverty reduction, it is not sufficient. Not all development policies are pro-poor; some are anti-poor. Policies like financial and capital market liberalization have, at least in some countries, contributed to greater instability, and a consequence of that instability is more poverty. Contractionary monetary and fiscal policies in response to crises exacerbate the downturns, leading to higher unemployment and a higher incidence of poverty. Policies to promote a learning economy too can either be pro- or anti-poor, but the most successful policies will necessarily be broad-based, engendering a transformation of the learning capacities of all citizens, and will therefore be pro-poor.

The second comment relates to the broader objectives of development, which should be sustainable improvements in the well-being of the citizens of the country, and the metrics we use to assess success. Our metrics don’t typically capture the increase in the wealth of a country that is a result of the learning strategies advocated here. It is only gradually, over time, that the benefits are realized and recognized.

The aftermath of the global financial crisis should be an exciting time for economists, including development economists, since it dramatically revealed flaws in the reigning paradigm. This paradigm has had enormous influence in development economics, though that influence was already waning, because its prescriptions had failed. Fortunately there are alternative frameworks already available—a plethora of ideas that should provide the basis for new understandings of why a few countries have succeeded so well and some have failed so miserably. Out of this understanding, perhaps we will be able to mold new policy frameworks that will provide the basis of a new era of growth—growth that will be both sustainable and enhance the well-being of most citizens in the poorest countries of the world.

Notes

This article was originally a paper prepared for a World Bank symposium, based on Justin Yifu Lin’s paper, “New Structural Economics.” The perspective taken here is based on joint work with Bruce Greenwald (2006 forthcoming). I am indebted to Eamon Kirchen-Allen for research assistance.

1. Even before Solow, Schumpeter had argued that the strength of a market economy resided in its ability to promote innovation and invention; and, shortly after Solow’s work, there developed a
large literature on endogenous growth, associated with names like Arrow, Shell, Nordhaus, Atkinson, Dasgupta, Uzawa, Kennedy, Fellner, and Stiglitz, followed on in the 1980s and 90s by the work of Romer. (See for example Atkinson and Stiglitz, 1969; Dasgupta and Stiglitz, 1980a and 1980b; Fellner, 1961; Kennedy, 1964; Nordhaus, 1969a and 1969b; Romer, 1994; Shell, 1966 and 1967; Uzawa, 1965.) The earlier work on endogenous (sometimes referred to as induced) innovation addressed not only the rate of innovation but its direction. For a discussion of more recent contributions in this line of research, see Stiglitz (2006).

2. Indeed, the work of Krugman has emphasized that today most trade is not related in fact to differences in factor endowments.

3. I use the term broadly to embrace any policy attempting to affect the direction of the economy.

4. Indeed, if all projects were successful, it suggests that the government is undertaking too little risk.

5. As I have also noted, such policies may have an adverse effect in enhancing domestic learning capacities.

6. The International Commission on the Measurement of Economic Performance and Social Progress emphasized the failures of GDP to reflect either sustainability or well-being (Fitoussi, Sen, and Stiglitz 2010). GDP per capita does not say anything about how well most citizens are doing; it can be going up even though most citizens incomes are declining (as has been happening in the United States). GDP focuses on production in the country; not on incomes earned by those in the country, and takes no account of environmental degradation or resource depletion, or, more broadly, of sustainability. The United States and Argentina both provide examples of countries whose growth appeared to be good—but both were based on unsustainable debts, used to finance consumption booms, not investment.

References


Gender and the Business Environment for New Firm Creation

Leora F. Klapper • Simon C. Parker

The authors summarize the extant literature on the relationship between gender and entrepreneurship. They note significant quantitative gender differences in business entry, with male-owned firms heavily prevailing over firms owned by women in many parts of the world. They find that enterprises owned by men on the one hand and women on the other are generally concentrated in different sectors, women entrepreneurs being better represented in labor intensive sectors such as trade and services rather than capital intensive manufacturing industries. They also observe certain gender differentials in business survival and growth patterns. Yet an analysis of a large body of literature does not suggest that, in general, the so called “gender gap” in entrepreneurship can be explained by explicit discrimination in laws or regulations. Rather, differences in quantitative and qualitative indicators of business entry and performance can in part be explained by a number of business environment factors that disproportionately affect a woman’s decision to operate a business in the formal sector. For example the concentration of women in low capital intensive industries—which require less funding and at the same time have a lower potential for growth and development—might also be driven by barriers against women regarding access to finance. Furthermore, women may have relatively less physical and “reputational” collateral than men, which limits their access to finance. Overall the literature suggests that improvements in the business environment can help promote high-growth female entrepreneurship. JEL codes: L26, J16, H11, O16, O17, P43

Increasingly policymakers are exploring ways of promoting economic activity and growth among women in developing countries. With the success of Grameen and other microfinance schemes which lend mainly to women, it is becoming apparent that female entrepreneurship represents a potentially valuable tool for
promoting growth and reducing poverty. Yet relatively little is known about the role of female entrepreneurship, especially in developing countries, or about the opportunities and barriers that women entrepreneurs face in practice.

The aim of this article is to review the literature relating to these questions and to speculate about the underlying reasons for observed findings. We start in the next section by outlining some salient facts about the extent of formal female entrepreneurship and the performance of women entrepreneurs. It is seen that although the ratio of female to male entrepreneurship varies across countries (Reynolds, Bygrave, and Autio 2004; Bosma and Harding 2007; Parker 2009, ch. 6), women entrepreneurs are outnumbered by men in many parts of the world—often by a wide margin (Estrin and Mickiewicz 2009).

We then describe how women tend to lag behind men in terms of the turnover, profitability, and growth performance of their businesses. These findings are well established in both industrialized and developing countries and motivate our primary question of whether these outcomes reflect gender differences in voluntary individual choices, or constraints in the business environment that restrict the opportunities of female entrepreneurs more than male ones.

We find some evidence that in developing countries female entrepreneurs might be constrained by features of the investment climate, such as barriers against access to credit and fair legal treatment. These issues are discussed and relate to questions such as gender discrimination, the role of property rights, and gender-based social norms.

Throughout the paper, we address the recurring question of whether individual choices or the business environment are primarily responsible for lower rates of female entrepreneurship and their lower average financial performance. The arguments behind these positions inform our policy recommendations, which we summarize.

**Women’s Engagement in Entrepreneurship**

This section commences with some stylized facts about the nature and extent of women’s participation in entrepreneurship. We then move on to discuss briefly three particular issues: household factors, industry choices, and motivations, both intrinsic and extrinsic. The section closes with us asking whether the different rates of participation in entrepreneurship primarily reflect choices or constraints.

An important caveat that should be mentioned is that our analysis, in general, focuses on entrepreneurship in the formal private sector. However, the informal sector (or “shadow economy”) plays an important role in many countries, ranging from over 75 percent of official GDP in Nigeria to about 10 percent in the United States (Schneider and Enste 2000). The previous literature has
highlighted the potential advantages to formal sector participation, including police and judicial protection (hence less vulnerability to corruption and the demand for bribes), access to formal credit institutions, the ability to use formal labor contracts, and greater access to foreign markets (Schneider and Enste 2000). However, because of burdensome regulations, high marginal tax rates, the absence of monitoring and compliance (of both registration and tax regulations), and other weaknesses in the business environment, many firms might find it optimal to evade regulations and operate in the informal sector. Firms that choose to stay small and informal might be unable to realize their full growth potential. Our discussion spotlights barriers in the business environment that challenge female entrepreneurs in the formal sector, who have the greatest potential for job creation and high-growth entrepreneurship.

The basic facts about women’s rates of participation in entrepreneurship are stark. Regardless of whether ‘entrepreneurship’ is defined in terms of ‘new venture creation’, ‘business ownership,’ or ‘self-employment,’ a higher proportion of men than women engage in this activity in industrialized economies. For example, according to data from the US Census Bureau, there were 6.5 million privately held women-owned firms in the United States in 2002, accounting for just over one-quarter of all firms. This number was up by 20 percent over 1997–2002: twice the growth rate of U.S. firms as a whole (Robb and Coleman 2009). Women entrepreneurs are also in the minority in Europe, where female self-employment rates in the EU vary considerably, from just over 20 percent in the UK, Ireland, and Sweden to 40 percent in Belgium and Portugal (Cowling 2000). Data collected from the business registrars of companies in Azerbaijan, Italy, Romania, and Tajikistan reveal that the ratio of new female-owned firms (as a percentage of total new firms) varies from 8.5 percent in Tajikistan to 38.3 percent in Romania. In Azerbaijan and Italy the ratio is 23.4 and 31.1 percent, respectively. Overall the data suggest that women entrepreneurs are even less represented in developing countries. Further evidence using 2005 firm level data for 26 countries in Eastern Europe and Central Asia finds that the average share of female entrepreneurs in the region is 28 percent, varying from over 40 percent in Latvia and Hungary to less than 15 percent in Armenia and Albania (Sabarwal and Terrell 2008). In addition this study finds that the majority of firms with less than 10 employees are female owned, while the reverse is true for larger firms. Furthermore female-owned firms in Eastern Europe and Central Asia constitute nearly 70 percent of “unspecified service” firms, in comparison to less than 10 percent of mining and construction firms.

Women entrepreneurs in industrialized countries are also more likely to be part-time workers than (i) women employees are and (ii) men are, in either employment category. For example calculations using data from the British Household Panel Survey (ISER 2008) reveal that females comprise only 16
percent of full-time employees, but as much as 70 percent of the part-time, self-employed workforce in the United Kingdom (Parker 2009, ch. 6). In the United States there appears to be two distinct groups of women entrepreneurs, with one group working less than 15 hours per week in their business and the other working full-time (more than 40 hours per week). Budig (2006) argues that the first group of women engage in nonprofessional self-employment primarily to limit their work hours and juggle family commitments (possibly because nonprofessional waged jobs tend to be less family friendly), whereas the second group enters professional self-employment to advance their careers. These arguments are based on evidence that family factors, especially children, help to explain the entrance of women into nonprofessional self-employment, though (as in the case of males) they have little impact on female entry into professional and managerial self-employment. Budig concludes that “women entering self-employment in professional occupations are more similar to their male peers in self-employment than they are to women entering non-professional self-employment” (2006, p. 2235).

It is certainly well established that men systematically contribute less to household production than women, even when the woman is working, and whether or not they are engaged in paid-employment or entrepreneurship. This finding is based on time-use data and responses to household surveys (Longstreth, Stafford, and Mauldin 1987; Boden 1999; Bond and Sales 2001). With less time to spend on formal work, part-time entrepreneurship can offer married women the flexibility to combine home and work commitments. This might partly explain higher female rates of part-time participation in entrepreneurship. Being married and raising children are both strongly associated with self-employment among women (Parker 2009, ch. 6). Having children of less than six years of age has the greatest impact on the probability that women are self-employed, especially among homeworkers (Edwards and Field-Hendrey 2002).

A 2005 Eurostat survey of entrepreneurs across 15 EU countries on start-up motivations also finds a significant gender difference in the decision to allocate more time to family life. For instance while only 7 percent of male respondents mention the relative maturity of their children as an incentive to start a business, this motive was highlighted by 16 percent of female respondents. Women are also more likely to cite the ability to combine work and private life—30 percent of women versus 42 percent of men—as the motivation to start up their own firms.

Another important gender difference in entrepreneurial participation is the industries in which businesses are established (Verheul, Risseeuw, and Bartelise 2002; Greene and others 2003). Women entrepreneurs remain heavily over-represented in a few industry sectors, especially sales, retail, and services. For example fully 69 percent of women-owned firms were in the service sector in 2006, while
14.4 percent were in retail trade (CWBR 2008). Similar sectoral distribution patterns, as well as a high concentration of female-owned firms in low-income informal sectors, have also been found in developing countries such as Indonesia (Singh, Reynolds, and Muhammad 2001). In contrast only a small percentage of women-owned businesses are located in high-growth or high-technology sectors (Menzies, Diochon, and Gasse 2004; Morris and others 2006).

Industry concentration is important because, as we will see in the next section, it has implications for the performance of women-owned ventures. The sectors that women cluster in are typically characterized by smaller scale, more intense competition and lower average returns. A question that naturally arises is whether industry clustering by women entrepreneurs reflects choices or constraints in the business environment.

To the extent that entrepreneurs identify opportunities to start businesses of similar types and in similar industries in which they formerly worked, one might be able to explain a portion of the industry concentration of women entrepreneurs in terms of different labor market experiences that vary by gender (Carter, Williams, and Reynolds 1997). For example female wage-and-salary workers are heavily concentrated in clerical and administrative jobs which require less advanced qualifications and which yield work experience that is arguably ill-suited to switching into entrepreneurship (Boden 1996). Male entrepreneurs in contrast are more likely than females to have been employed prior to start-up. They also have more previous work and business experience in industry and in managerial roles on average (Brush 1992; Carter, Williams, and Reynolds 1997; Boden and Nucci 2000; Kepler and Shane 2007; Fairlie and Robb 2009). And male entrepreneurs are more likely than females to have education and experience with technical, business, or managerial elements (Brush 1992; Menzies, Diochon, and Gasse 2004). In order to prepare females for a broader range of industry choices, including nontraditional fields such as engineering and science which offer higher growth potential and greater rewards, it has been argued that they should be encouraged to study business and technical subjects (Hisrich and Brush 1984, 1987; Hisrich 1989).

Gender differences in women’s choices of industry and the sizes of their businesses might alternatively be linked to differences in business objectives. It has been suggested, for example, that women are less motivated than men by growth and profits, and more by internal goals such as personal fulfillment, flexibility, and autonomy (Anna and others 1999; Carter and others 2003; Morris and others 2006). However, several empirical studies of both nascent (potential) and actual entrepreneurs do not find any significant gender differences in terms of (self-declared) growth or profit motives for business start-up. Similar to male entrepreneurs, the drive of women for personal freedom, independence, satisfaction, and/or security as motivating factors for running a business is mentioned in
studies conducted in countries such as the United States, Singapore, Norway, and Pakistan (Hisrich and Brush 1985; Ljunggren and Kolvereid 1996; Shabbir and Di Gregorio 1996; Maysami and Goby 1999; Scheiner and others 2007). And a recent German study found that entrepreneurial self-perception was an equally important motivator at business start-up for males and females (Werner and Kay 2006).

More clear-cut results emerge when one considers the role of external factors. For instance a study conducted in Norway found that women are more prone than men to be influenced in their decisions regarding starting a business by external influences like family or community opinions (Ljunggren and Kolvereid 1996). And in Italy, whereas men tend to enter self-employment for career advancement considerations, women tend to enter from inactivity or unemployment (Rosti and Chelli 2005). Furthermore a novel dataset collected during a natural experiment in a rural setting in the Appalachian region of the United States in the post-Tobacco Buyout era supports the hypothesis that females, but not males, are pushed into entrepreneurial activities by changing economic environments and a lack of household income (Pushkarskaya and Marshall 2008). This is consistent with the 2005 Business Success Survey conducted by Eurostat in 15 European countries which found that women are significantly more likely to report their motivation for starting their own business “to avoid unemployment” than men (58 percent of women versus 42 percent of men). Overall then the literature suggests that female entrepreneurs’ motivation is relatively more likely to represent a job transition or a re-entry into the workforce following a lay-off or voluntary leave (Kaplan 1988).

So do gender differences in entry rates and sectors reflect gender-based differences in choices and attitudes, or do they stem from differences in the external business environment? While there can never be a definitive answer to this question, the evidence cited above suggests that choices of career paths, work experience, and fertility all play important roles, as do motivations deriving from business conditions and prospects in the broader economy. This would seem to support the notion that women consciously choose their entrepreneurial engagement profiles. On the other hand it could be argued that women are socialized into taking these choices and that the socialization process inculcates and replicates dominant patriarchal norms. While we lack hard evidence on this proposition, anecdotal evidence suggests that the social norms argument may have some relevance in some developing countries in which overt restrictions on women’s activities are observed. But even in industrialized countries, the ongoing gender differential in terms of the burden of household chores, which continues to fall largely on women, raises deep questions about underlying gender roles. Unfortunately we still do not know enough about the relative importance of subtle yet persistent social norms...
(“nurture”) or the role of biology in shaping tastes and predispositions (“nature”). By comparison more is known about explicit barriers to female entrepreneurship in the form of access to finance, an issue which will be discussed in some depth below.

**Women’s Performance in Entrepreneurship: Survival, Growth, and Profitability**

This section commences with a brief review of evidence about women’s performance in entrepreneurship. This can be measured in several ways, including earnings and profits, rates of return on capital, venture growth rates, and survival. A key finding that emerges is that in terms of all of these metrics, women entrepreneurs tend to underperform relative to their male counterparts. We go on to consider several possible explanations for this finding and then discuss whether gender differences in performance reflect choices or constraints.

There is broad agreement among researchers that, in both industrialized and developing countries, women entrepreneurs earn less income than male entrepreneurs. For example measuring entrepreneurship in terms of self-employment, full-time self-employed American women in 1990 earned 73 percent of the annual income of women wage-and-salary workers, whereas self-employed men earned 107 percent of the annual income of their employee counterparts (Devine 1994). In this study self-employed men were more frequently observed to work in high-paying sectors than women, including executive, administrative, managerial, and precision artisan jobs. In contrast self-employed women were more frequently found in lower-paying service and retail sectors. These sectors are more competitive and also exhibit lower business survival rates.

Self-employed women perform somewhat better on average in other countries, with female/male earnings ratios reaching 87 percent in the case of Australia (OECD 1986). But in developing countries women receive substantially lower average returns than men do, especially if they are not employers (for example see Honig 1996, 1998, for Jamaican evidence).

A similar story applies if performance is measured in terms of turnover, employment creation performance, growth rates, or survival prospects. For example Loscocco and Robinson (1991) report that women generate only one-quarter of men’s average business receipts, while Bosma and others (2004) estimate that male Dutch business founders outperform their female counterparts in terms of survival, profitability, and employment creation. With regard to growth, one longitudinal study found that the majority of female-owned businesses were moderately successful with revenue increases of approximately 7 percent per year, though this was significantly less than the average for male-owned firms (Hisrich
and Brush 1987). And for developing countries, a recent study conducted in Sri Lanka reported that female microenterprise owners had returns to capital that were dramatically lower than male entrepreneurs—in many instances zero or even negative (de Mel, McKenzie, and Woodruff 2009).

It should be acknowledged that the evidence on inferior women entrepreneurial performance is not unanimous. It is true that some studies have found similar survival and growth rates for male- and female-owned businesses (Kalleberg and Leicht 1991; Westhead, Storey, and Cowling 1995). But these studies are the exception rather than the rule.

What explains these findings of inferior female-owned business performance? As observed in the previous section, women entrepreneurs tend to be concentrated in industries with lower capital intensities and lower average returns to capital. Some researchers have attributed women entrepreneurs’ inferior financial performance to this factor (Loscocco and Robinson 1991; Loscocco and others 1991; Hundley 2001). Another common observation is that women entrepreneurs operate smaller businesses on average, utilizing less capital and finance from banks and other lenders than men do (Aronson 1991; Carter, Williams, and Reynolds 1997; Watson 2002). For example Watson’s (2002) analysis of Australian Federal Government data on 13,551 male small and medium sized (SME) business owners and 875 female SME business owners revealed that the females earn similar rates of return on equity and assets as males, though the former invested less to start with, which explains why they ended up with lower absolute income and profits than the males.

It is unclear whether these factors reflect different preferences or different responses to external business conditions. Some insight into this issue can be gleaned by looking at the self-reported preferences and goals of women entrepreneurs. Studies conducted in Canada have shown that running a “small and stable business” was a preferred mode of operation among women but not among men (Lee-Gosselin and Grise 1990; Cliff 1998; Verheul, Thurik, and Grilo 2008). And according to another study, self-employed female accountants report lower profitability of their firms to be an acceptable result of their choice for greater flexibility and to be able to devote more time to their families. At the same time higher profitability projections were associated with higher achievement and income goals (Fasci and Valdez 1998).

It might also be thought that greater risk aversion among women than men might explain their lower average returns and growth performance in entrepreneurship, leading women to choose positions further down the expected profit–risk frontier than their male counterparts (Watson and Robinson 2003). Although Watson and Robinson (2003) observed that female-owned businesses in Australia were characterized by both lower risk and lower returns, with similar risk-adjusted returns by gender, the evidence about gender differences in risk
attitudes is sparse and inconclusive. While Jianakoplos and Bernasek (1998) adduced survey evidence suggesting that women are more risk averse than men, experimental evidence (albeit from a small sample of students) from Schubert and others (1999) detected similar levels of risk aversion among men and women.

Similar to the findings about women entrepreneurial participation rates discussed in the previous section, preferences that are associated with household production seem to play an important role. A decomposition analysis by Hundley (2001), which explored the role of numerous explanatory variables, is especially instructive. According to Hundley the tendency of women to do more housework, work fewer hours in business, and do more childcare accounted for between 30 and 50 percent of the American annual self-employment gender earnings differential. This suggests that women earn less than men do because they spend more time in household production and less time managing and developing their businesses.

Complementary evidence comes from a study which reports that rather than reinvesting their profits, women entrepreneurs in Morocco are more likely to spend their income on family and household needs, save cash for emergencies, or both (Murray and Barkallil 2006). Furthermore a field experiment of providing random grants to microenterprise owners in Sri Lanka finds that grants generated large profit increases for male entrepreneurs, but not for females. The authors report that this can in part be explained by inefficient resource allocation within households, which is reduced in more cooperative ones (de Mel, McKenzie, and Woodruff 2009).

Yet external constraints also seem to play an important role in transition and developing countries. Thus studies of formal enterprises in Eastern Europe and Central Asia find that women-owned businesses tend to be small and use less external financing (Sabarwal and Terrell 2008). Sabarwal and Terrell (2008) also find that returns-to-scale of women-owned firms are significantly larger than male-owned firms, implying that women entrepreneurs would gain more from enlarging their firms, relative to male entrepreneurs. This study finds that a main reason for the suboptimal size of women-owned firms is capital constraints. Similar results have been found in Latin America and Africa (Sabarwal and Terrell 2008).

The current state of the literature does not clearly reveal whether gender performance differences in entrepreneurship are primarily attributable to voluntary choices by women entrepreneurs or to constraints. Yet the evidence cited above provides some interesting clues. Suppose one identifies different levels of household production with preferences, and different access to finance with external constraints. Then the findings of Hundley (2001) and Sabarwal and Terrell (2008) might lead us to attribute performance differences in industrialized countries mainly to preferences (especially those related to child rearing and Klapper and Parker 245
Challenges in the Business Environment and Policy Implications

In the previous sections we documented differences between male and female entrepreneurs—in business participation rates, sectoral choices, growth rates, profitability, and survival rates—and discussed the competing hypotheses of whether these differences are driven by choices or barriers in the business environment. In this section we review possible institutional barriers to female entrepreneurship, including credit constraints, property rules, and adverse social norms. Evidence from cross-country studies shows that women receive a lower share of available external funding than men for business and other purposes. Yet the evidence suggests that this might be less attributable to explicit discrimination than to weaknesses in the business environment that make lending to women a higher (perceived or actual) credit risk.

An extensive literature now documents the positive effect of financial development on economic growth in both industrialized and developing countries (Rajan and Zingales 1998; Levine, Loayza, and Beck 2000). A related literature demonstrates a robust relationship (using data on over 100 countries) between firms’ access to finance and various institutional features of the business environment, including creditor rights (La Porta and others 1998), the credit information infrastructure (Djankov, McLiesh, and Shleifer 2007), and bankruptcy regimes (Djankov and others 2008). Studies have also found that financial development has an impact on new business creation (Klapper, Laeven, and Rajan 2006). These papers highlight the importance for entrepreneurs of access to external financing, particularly for the formation of larger, capital intensive firms.

Yet cross-country studies have shown that women are less likely to get financing from a formal institution or are charged a higher interest rate than men (Muravyev, Schäfer, and Talavera 2007; Demirgüç-Kunt, Beck, and Honohan 2008). Similarly women entrepreneurs generally raise less formal and informal venture capital than men (Brush and others 2004). However, other evidence from industrialized countries, such as the United Kingdom, the United States, Canada, and New Zealand, and from certain developing countries, such as Ecuador and Peru, shows that women applying for funding generally do not face arbitrarily higher denial rates than men, suggesting that gender differences in the use of credit might be explained by differences in the demand for external financing (Buvinic and Berger 1990; Aguilera-Alfred, Baydas, and Meyer 1994; Baydas,
Meyer, and Aguilera-Alfred 1994; Coleman 2000, 2002; Carter and Shaw 2006). In particular fewer women than men apply for funding (Buvinic and Berger 1990) and women generally request smaller amounts (Aguilera-Alfred, Baydas, and Meyer 1994; Coleman 2000).

In many developing countries (including Bangladesh, Malawi, India, Pakistan, Ethiopia, Kenya, Tanzania, Uganda, and Zambia) and transition economies, women entrepreneurs report facing greater and more systemic access barriers to formal financial services; and they cite finance as a major challenge in starting and growing their businesses (Rose 1992; Diagne, Zeller, and Sharma 2000; Goheer 2003; Faisel 2004; ILO/AfDB 2004; Richardson, Howarth, and Finnegan 2004; GEM/IFC 2005; Bardasi, Blackden, and Guzman 2007; Ellis and others 2007a, 2007b; Demirgüç-Kunt, Beck, and Honohan 2008; Narain 2009). Further evidence from Asia and Africa finds that women entrepreneurs are more likely than men to rely on internal or expensive informal financing when formal funding is unavailable (Rose 1992; Richardson, Howarth, and Finnegan 2004).

Importantly these studies do not find explicit discrimination against female borrowers. It is therefore important to identify why women might face greater challenges than men in accessing formal finance. As discussed in previous sections, female entrepreneurs might choose to enter less capital intensive industries. The disinclination of women to seek external financing might also be attributed to their own perception that they will have a harder time securing bank loans (Coleman 2000).²

Alternatively, female loan applicants, relative to male applicants, might have weaker loan applications. For instance women are found to have relatively greater difficulties completing complicated loan applications (Buvinic and Berger 1990). They also have lower financial literacy (Lusardi and Tufano 2009), which might make it harder for them to navigate the loan market.

In addition, and as noted already, women have been found to have weaker business backgrounds than men, including a lack of relevant education (especially technical) and a lack of business experience (Carter and others 2003; Menzies, Diochon, and Gasse 2004). Women are also less able on average to provide collateral (Buvinic and Berger 1990) or personal guarantees (Coleman 2002). Female entrepreneurs might also have weaker credit histories (“reputational collateral”), because loans, utilities, cell phones, and other debts might be in their husband’s name. These findings suggest that women, on average, might have lower credit scores, which are important for modern lending technologies (Narain 2009).

A more recent strand of literature cites “behavioral” patterns on behalf of lenders and borrowers. Studies of online lenders have shown that photographs of borrowers affect loan decisions and loan terms for both bank loans and microfinance (Ravina 2008; Pope and Sydnor 2008; de Laat and Chemin 2008). For instance a study of about 12,000 loan applications finds that a higher “beauty
rating” of the borrower increases the probability of getting a loan by 2.04 percent for women, but only 0.6 percent for men (Ravina 2008).

However, these findings only invite further explanation. For instance why do women have relatively less access to physical or reputational collateral than men? And what policies can address these barriers to access to finance, such as changes in laws or behavior that would allow women to acquire the collateral necessary for start-up financing?

One explanation is weaknesses in the legal environment, which can affect the ability of lenders to collateralize assets, and to seize them in the case of default. Although there are few examples of business and economic laws that explicitly discriminate against women, other rules and regulations may hamper female entrepreneurship. For example the requirement that married women in many Middle Eastern countries receive the permission of their husbands to obtain a passport or to travel can be a significant impediment to doing business (Chamlou 2008). Property laws are also important: married women may not be deemed creditworthy since they do not possess the title to their land or house, which might be due to unequal laws of inheritance (Goheer 2003; ILO/ADB 2004; GEM/IFC 2005; Ellis and others 2007b; Morrison, Raju, and Sinha 2007; Demirgüç-Kunt, Beck, and Honohan 2008).

Other laws might disproportionally restrict female, relative to male, entrepreneurs. These include requirements for married women to obtain their husband’s signature and approval for all banking transactions (South Africa and Uganda). Women can also be affected by a husband’s adverse credit history, which might require his wife to repay the debt or be denied credit (South Africa) (Blanchard, Zhao, and Yinger 2005; Naidoo and Hilton 2006). Although men may also have to repay their wife’s debt under the same circumstances, it is more likely that the husband has incurred previous debts.

Social norms and differential treatment under the law are also important. Although women might not be legally prohibited from obtaining licenses required for accessing certain types of financing, evidence from Africa shows that in many instances only male heads of households are able to receive them successfully (Johnson 2004; Narain 2009). Similarly, in many Middle Eastern and South Asian countries, women are required in practice to have a father or husband co-sign a loan, even though banking laws do not require it in principle (Chamlou 2008).

To address these institutional barriers in developing countries, various government and institutional policies have been tried in an effort to increase the number of women borrowers. These include financial literacy training programs; public awareness workshops; business development services (Bangladesh and South Africa); promotions of sectors dominated by women entrepreneurs (Bangladesh, India, West Africa, Chile, the Philippines, and Canada); marketing
support; business training (Peru, India); legal aid; and female business networks. An important caveat is that few of these studies included randomized or control samples, which are necessary for ascertaining that women would not have increased their borrowing without the benefit of a government program.

Other improvements in the business environment might also have a disproportionate impact on female entrepreneurs. For instance introducing credit registries and bureaus for consumer and commercial borrowing can help both men and women build credit histories, although this might have a relatively greater impact on women who are less likely to have physical collateral in their names. In addition women are likely to benefit relatively more than men from the inclusion in credit registries of nonbank credit information, such as utility payments, cellphone histories, and consumer finance transactions.

Women might also benefit from the development of alternatives to physical collateral, which could allow banks to make less risky loans to women who are unable to pledge collateral. This includes co-signature loans, using savings and remittance behavior as a measure of determining credit worthiness, and incentive schemes such as conditioning new loans upon successful repayments (Almeyda 1996).

Another important measure would be to reform collateral laws that cover movable property. That might overcome women’s lack of control over immovable assets in many developing countries (Fleisig, Safavian, and de la Pena 2006). Promoting a tax, legal, and regulatory environment that promotes business equipment leasing might also be a particularly important financial product for women who do not have land to use as collateral, have no banking histories, or who have limited start-up capital (Dowla 2000; Bass and Henderson 2000; Ellis and others 2007b; Brown, Chavis, and Klapper 2008).

Conclusion

As reviewed there is a substantial literature in industrialized countries, and a growing literature in developing and transition countries, that uses microlevel data to study women’s entrepreneurship. Previous studies primarily focus on gender differences in business intentionality, human capital and access to finance, and the impact of these factors on business entry and performance. Women’s educational backgrounds in the social sciences rather than technical disciplines, and their desire to combine entrepreneurship and family work, are often offered as explanations as to why women entrepreneurs are overwhelmingly concentrated in highly competitive, small-scale, labor intensive businesses. In addition similar gender-based sectoral concentration patterns, as well as a large representation of female entrepreneurs in informal segments of the economy, have been observed in
developing countries. However, the concentration of women in low-capital intensive industries—which require less funding and at the same time have a lower potential for growth and development—might also be driven by women’s barriers to access to finance.

Few studies conducted in either industrialized or developing countries have analyzed gender implications of the business regulatory environment. There are still many outstanding questions, such as the impact of improvements in the credit information infrastructure on women’s ability to mobilize resources for business creation and development. Or the impact of greater access to start-up capital on the industry-selection of female entrepreneurs. Addressing these questions can help policymakers target reforms to promote high-growth female entrepreneurship.

Notes

Leora F. Klapper is Senior Economist in the Development Research Group, the World Bank, 1818 H St., NW Washington, DC 20433, USA; tel.: 1-202-473-8738; email address: lklapper@worldbank.org. Simon C. Parker is Associate Professor at the Richard Ivey School of Business, University of Western Ontario, 1151 Richmond St N, London, ON N6A 3K7, Canada; tel.: 1-519-661-3861; email address: sparker@ivey.uwo.ca. The authors thank Anna Cusnir, Amanda Ellis, David McKenzie, Juan Manuel Quesada Delgado, Ana Ribeiro, and three anonymous referees for valuable comments. The opinions expressed do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.


2. In addition to the reasons listed above, another explanation might be female under-representation among bankers and venture capitalists (Aspray and McGrath Cohoon 2006); however, these studies do not find any evidence of explicit discrimination among male lenders and investors.


4. However, previous studies have found that the introduction of credit information systems can lead to the initial screening out of poor individuals, which include a disproportionate number of women borrowers (McIntosh, Sadoulet, and de Janvry 2006; Malhotra and others 2006; Luoto, McIntosh, and Wydick 2007).

References

The word processed describes informally reproduced works that may not be commonly available through libraries.


Dessy, Sylvain, and Jacques Ewoudou. 2006. “Microfinance and Female Empowerment.” Cahiers de recherche 0603, CIRPEE.


*Klapper and Parker* 255


Explaining Enterprise Performance in Developing Countries with Business Climate Survey Data

Jean-Jacques Dethier • Maximilian Hirn • Stéphane Straub

The authors survey the recent literature which examines the impact of the business climate on productivity and growth in developing countries using enterprise surveys. Comparable enterprise surveys today cover more than 100,000 firms in 123 countries. The literature that has analyzed this data provides evidence that a good business climate favors growth by encouraging investment and higher productivity. Various infrastructure, finance, security, competition, and regulation variables have been shown to have a significant impact on enterprise performance. The authors state their motivation for their review by explaining why a disaggregated, firm-level analysis of the relationship between enterprise performance and business climate—as opposed to a more macroaggregate analysis—is important to gaining insights into these issues. They review the main findings of the empirical microliterature based on enterprise surveys and consider the robustness of the results. To conclude they put forward some ideas to advance research on business climate and growth, and they suggest possible improvements in survey design. JEL codes: L51, O47, O12

In recent years, an unprecedented data collection effort has yielded a set of comparable enterprise surveys covering more than 100,000 firms from 123 countries. As a result, a number of studies have started to analyze the impact of the business climate variables contained in these surveys on different dimensions of firm performance. The general aim of this literature is to generate policy prescriptions based on the identification of the main constraints facing firms. Although many of these studies identify relevant constraints, contradictory or fragile results are also found, pointing to some weaknesses in the methodology.
applied in some papers as well as in the original survey questionnaire design itself.

We review the literature, highlight its main strengths and shortcomings, and propose potential improvements. We begin by stating our motivation for this review, explaining why firm-level analysis—as opposed to a more macroaggregate analysis—of the relationship between enterprise performance and business climate is important for gaining insights. We then review the empirical microeconometric literature using business climate survey data, and consider the robustness of the findings. Finally we put forward some ideas for advancing research on the business climate and suggest possible improvements in survey design. Appendix A presents a comparative view of the key studies analyzing the relationship between enterprise performance and business climate, focusing on datasets, methodology, main variables, and results obtained.

Economic Growth and the Business Climate

Many structural, institutional, and behavioral variables shape and drive economic growth. The critical variables that collectively define the business climate (also called investment climate) are infrastructure, access to finance, security (meaning the absence of corruption and crime), and the regulatory framework, including competition policies and the protection of property rights. The main hypothesis here is that the business climate affects economic activity throughout the economy and particularly through its influence on incentives to invest. An improvement in the business climate increases returns to current lines of activity and so increases investment in these. It also creates new opportunities—for example through trade or access to new technology. It influences the psychology of entrepreneurs—the Keynesian “animal spirits”—affecting their assessment of whether innovation will pay off. It puts competitive pressure on firms that have enjoyed privileged positions as a result of import or other protection, or special access to government officials. As a result of greater competition, it may cause some firms, perhaps those closer to technological frontiers, to succeed—even as others fail.

Given the complexity of effects that changes in the business climate elicit, different firms, industries, and regions will be affected in different ways. Moreover growth fueled by the business climate is not simply a shift toward some technological frontier. Developing countries must overcome or reduce all kinds of obstacles to efficiency, dynamic and otherwise, without any illusions that the economy will soon reach the frontier. Indeed changes in the business climate may have their most crucial impact far from the technological frontier.

A weak business climate, on the other hand, may not only discourage investment, it can also lead businesses to take costly or counterproductive steps to
defend themselves from the consequences of its weaknesses. If social order and
control are weak, firms typically have to invest heavily in defensive measures such
as private security. If the power supply is unreliable, firms will invest in backup
electricity generation capacity. If it is difficult to get goods through or to ports,
trade is discouraged and larger, more costly inventories are held. Many such con-
straints on development are not quickly or easily reversed.

On the contrary, improvements in the business climate could generate extra
growth dividends through political economy mechanisms if they increase the
number of people and enterprises with a stake in a better climate. For example, if
trade reforms create an export-oriented sector, that may increase pressure for
further reforms to trade policy or trade-related infrastructure. And higher
incomes might lead to pressure for an improved business climate in other ways,
as people seek rules governing the protection of wealth or capital.

There exists a rich macroeconometric literature which uses cross-country data
to relate broad indicators of institutional quality, policy, and infrastructure to a
number of macroeconomic outcome variables.¹ This has yielded interesting
insights—broadly speaking, that the business climate significantly affects eco-

On infrastructure, the seminal paper is Aschauer (1989), which finds that
infrastructure capital has a large impact on aggregate total factor productivity
(TFP). Many papers (reviewed in Straub 2008, 2010) since then have sought to
compare the elasticities of infrastructure capital and private capital. A number of
papers estimate a long-run aggregate production function relating GDP to phys-
ical capital, infrastructure (transport, power, and telecoms), and human capital.
A recent example—Calderón, Moral-Benito, and Servén (2010)—uses a dataset of
infrastructure stocks covering 88 countries and spanning the years 1960–2000
and finds statistically (and economically) significant estimates of the output con-
tribution of infrastructure. They imply, for instance, that an increase in infrastruc-
ture provision from the median lower-middle income country level (say, Bolivia in
2000) to that of the median upper-middle income country (say, Uruguay) would
yield an increase in output per worker of almost 5 percent. An increase in infra-
structure provision from median upper-middle income country level to median
high-income country level (say, Ireland) would raise output per worker by more
than 8 percent.

Regarding institutions and the policy environment, Pande and Udry (2005), as
well as Dollar, Hallward-Driemeier, and Mengistae (2005), find compelling evi-
dence that long-run growth is faster in countries that have higher quality legal
institutions, better law enforcement, increased protection of private property
rights, improved central government bureaucracy, smoother operating formal
sector financial markets, increased levels of democracy, and higher levels of trust.
World Bank (2004) finds that one of the useful insights of these macroanalyses
is that secure property rights and good governance are central to economic growth.

The findings of the macroeconometric literature are qualified by concerns about the robustness of the results. The possible endogeneity of infrastructure has been advanced as a reason for contradictory findings on the impact of public capital on long-run economic development. Endogeneity in this context might come from three sources: (a) measurement errors stemming from the use of public capital figures as proxies for infrastructure (see the discussion in Straub 2010); (b) omitted variables, which may arise when there is a third unobserved variable that affects the infrastructure and growth measure; and (c) the fact that infrastructure and productivity or output might be simultaneously determined, that is infrastructure provision itself positively responds to productivity gains.

The precise channels through which business climate variables affect economic growth are still not well understood, and recent studies have been more cautious in their interpretation of the evidence. Durlauf, Kourtellos, and Tan (2008) find some evidence that institutions play a role as determinants of GDP growth rates but they question the robustness of these results and state that the effect is likely to flow through the influence of institutions on physical capital accumulation rates and not via TFP growth directly. Straub, Vellutini, and Warlters (2008) find some evidence of a positive effect of infrastructure on growth, especially in poorer countries, but conclude that in general the results from studies using aggregate data lack robustness. Recent infrastructure elasticity estimates are much lower than earlier calculations which were often fraught with econometric problems such as not accounting for endogeneity or inefficient proxy variables (Romp and de Haan 2005). Other econometric problems, such as the failure to account for model uncertainty in cross-section studies, persist.

The macroeconometric approach has a number of inherent limitations:

- Results display considerable heterogeneity across economies. The explanatory variables at the country level obscure important dimensions of heterogeneity such as variations across different regions within a country, across different types of firms (by firm size, age, ownership type, etc.), or both.
- The limited number of countries restricts the sample size of country-level analyses, especially cross-sectional ones, and thus the robustness of the results.
- Aggregate business climate indicators are often imprecise, rely on de jure information, or subjective judgments about the relative weight of variable components, and lack direct input about actual conditions as experienced by affected parties such as firms.
- Most country-level indicators are invariant over time and thus are indistinguishable from fixed (country, sector, or region specific) effects that may reflect features other than the business climate (Commander and Svejnar 2007).
• The instruments most often used consist of geographical or historical preconditions (latitude, colonial history, settler mortality, etc.), which limits the ability of the empirical models to identify the consequences of institutional change for growth.

These comments suggest that econometric analysis at a more disaggregated level (firm or industry level) is required to achieve more robust results and leads to more precise policy recommendations—a point that is repeatedly emphasized in World Bank (2004), Pande and Udry (2005), and Durlauf, Kourtellos, and Tan (2008). Such a model should account for the behavior of individual firms in a world where either markets or governments fail, or people face psychological difficulties in taking advantage of opportunities, as suggested by Banerjee and Duflo (2005). To the extent that changes in the business climate affect different firms differently, an aggregate model of business climate and economic growth, with its reliance on a representative firm, is therefore inadequate. Indeed the heterogeneity of firms’ responses to changes in the business climate is likely to generate changes in their geographical, industrial, and size distributions.

Relatively simple disaggregated models, addressing the constraints of interest and taking enterprises characteristics (size, ownership, location, type of activities, etc.) into account, can provide a variety of insights; and empirical studies that exploit enterprise surveys are examples of the added value provided by a more disaggregated, microeconomic approach. Note that in the standard theory of profit-maximizing firms, prices of inputs are set equal to their marginal products, so all inputs should be equally “constraining.” The discussion of firm-level constraints (high prices of given inputs) would then involve deciding whether these are just intrinsic characteristic of the natural environment (for example the climate) or whether they derive from government failures. However, in practice some key inputs are often not priced at their marginal costs. For example, despite the increasing market mediation of infrastructure, there is also strong evidence that firms’ costs and prices are largely not reflecting the “fundamentals” of these activities, so it is implausible that this type of capital is remunerated according to its marginal productivity. In practice, this makes disentangling the sources of these constraints more difficult.

Recent Enterprise-level Business Climate Studies

Firm-level business climate data have proved to be a rich source of information on the characteristics of firms and the constraints they face in the developing and transitioning world. This section first describes the characteristics of these surveys and their evolution over time, then reviews the findings of the empirical literature that exploits these surveys to explain firm performance as a function of different aspects of the business climate.
Datasets

This section provides an overview of enterprise survey datasets, then examines the structure and content of the standardized core survey instrument. In the paper’s conclusions, we make some suggestions to improve the design of the questionnaire.

Overview of Existing Datasets. The crucial prerequisite for finding more disaggregated evidence is the availability of raw disaggregated data. Before the 1990s, standardized firm-level business surveys spanning multiple countries were practically nonexistent. This began to change with an initial series of largely self-contained projects which carried out business surveys for certain sets of countries and with various thematic scopes.

Four key projects of that period were sponsored by the World Bank: A first set of surveys carried out from 1992 to 1995 by the Africa Regional Program on Enterprise Development; the first round of the Business Environment and Enterprise Performance Survey (BEEPS) for 22 transition countries in 1999; the World Business Environment Surveys (WBES) implemented for 80 countries and the West Bank or Gaza territories from late 1998 to early 2000; and a number of Firm Analysis and Competitiveness Surveys by the Development Economics Research Group. While these projects yielded unprecedented and highly useful data for the countries and issues they were designed for, they suffered from limited comparability amongst each other due to differing questionnaire designs and priorities.

The key development of the early 2000s was a push for greater standardization in order to build up a single, centralized database of comparable business climate surveys from around the world. A set of core questions was “pooled and consolidated” from earlier surveys. This became the crucial component of the new, standardized business climate questionnaires known as Productivity and Business Climate Surveys (PICS). In a specific country survey, around 50–60 percent would consist of the core modules (some 80 questions), the rest would consist of nationally specific ones that could be added flexibly to the core instrument depending on each country’s data needs. The core instrument was also partly incorporated into the latest rounds of surveys that had started earlier, for instance BEEPS, the second and third round of which contain most of the core PICS questions. Launched in 2001, the new surveys have been used to acquire detailed firm-level data in 15 to 20 countries a year. The results have been collected in a central database (www.enterprisesurveys.org) along with those of earlier, comparable projects such as BEEPS II and III. All surveys in this database are now commonly referred to as Enterprise Surveys, although the old terminology (PICS, BEEPS, etc.) persists to some extent. By 2010, the database—which is accessible to anyone who registers—contains information on more than 100,000 firms in 123 countries. Aterido and Hallward-Driemeier (2007, p. 20) outline the key features of this database: “The median sample size is 350 firms, with several large countries having substantially larger samples… The sample of firms in each
country is stratified by size, sector and location. The unit of analysis is the ‘establishment’ in the manufacture and service sectors. Most firms are registered with local authorities, although they may be only in partial compliance with labor and tax authorities.”

The core questions are generally answered by the manager or owner of the establishment in face-to-face interviews. Accounting data may be provided by the establishment’s accountant, human resource manager, or both. Some countries have attached nationally specific modules answered by workers (for instance the Thailand 2007 survey—see World Bank 2008b). Among the earlier surveys, there is still some variation of the core questions, so that comparative analyses of multiple business climate variables may require a focus on a subset of the total database. But there is a large subset of firms and countries for which the data are comparable. Unfortunately important questions (about cost of electricity, number of power outages, or ownership of backup generators by enterprises, for instance) were dropped from most questionnaires after 2006, so that it is not always possible to compare all variables before and after that year.

**Structure and Content of the Core Business Climate Survey Instrument.** The standardized core survey instrument is organized into two distinct parts. The first part provides general information about the firm and the business climate it faces. The second part collects accounting information such as production costs, investment flows, balance sheet information, and workforce statistics. The questions about the firm and the business climate in the first part include:

- **General characteristics of the firm:** age, ownership, activities, location
- **Sales and supplies:** imports and exports, supply and demand conditions, competition
- **Business climate constraints:** evaluation of general obstacles
- **Infrastructure and services:** power, water, transport, computers, business services
- **Finance:** sources of finance, terms of finance, financial services, auditing, land ownership
- **Labor relations:** worker skills, status and training; skill availability; overemployment; unionization and strikes
- **Business–government relations:** quality of public services, consistency of policy and administration, customs processing, regulatory compliance costs (management time, delays, bribes), informality, capture
- **Conflict resolution or legal environment:** confidence in legal system, resolution of credit disputes
- **Crime:** security costs, cost of crimes, use and performance of police services
- **Capacity, innovation, learning:** utilization, new products, planning horizon, sources of technology, worker and management education and experience.
Both subjective data on perceptions of managers and objective data on various business climate indicators are recorded. Tables 1 and 2 are based on surveys covering 41,207 firms in 91 countries during the 2006–09 period. The first table ranks the perceptions of managers about issues that represent constraints for the operation of their enterprise, by geographic region. The severity of the constraints is also a function of the industrial branch they belong to, as shown in table 2.

There have been considerable discussions about the possible weaknesses of subjective, perception-based indicators compared to objective, quantitative data. Concerns have been raised whether subjective data may be vulnerable to waves of pessimism and euphoria, to inconsistencies across regions and countries because firms compare themselves to different benchmarks (so-called anchoring effects—Bertrand and Mullainathan 2001), or to managers’ inability to form accurate subjective estimates (Gelb and others 2007). For instance, managers may fail to separate internal weaknesses of the firm (for example inability to provide proper documentation) from external business climate constraints (for example inefficient bureaucracy). These problems are a specific concern when conducting econometric estimations based on cross-sectional data, and addressing them may require the use of panel data to control for individual or firm fixed effects.

Exploring such concerns, Gelb and others (2007, p. 30) examine subjective data yielded by the core Enterprise Survey perception questions. They conclude that while perceptions of critical business climate constraints may not always correspond fully to ‘objective’ reality, firms “do not complain indiscriminately,” and response “patterns correlate reasonably well with several other country-level indicators related to the business climate.” Aterido and Hallward-Driemeier (2007) underline that subjective rankings are highly correlated with objective measures in 16 of the 17 variables and also significantly correlated with external sources, including “Doing Business” indicators. Pierre and Scarpetta (2004), using data from 38 countries, confirm that countries with more restrictive labor regulations are associated with higher shares of firms reporting labor regulations as constraining.

Even if objective and subjective measures are significantly correlated, the latter remain prone to bias. For example, Olken (2009) compares corruption perceptions among villagers in Indonesia with objective measures of corruption in road construction projects. It shows that although these are positively correlated, there are also systematic individual-level biases in the latter. Similar issues are likely to arise in firm-level surveys.

In spite of these problems, subjective indicators can still play a useful role in identifying important constraints through descriptive statistics. Carlin, Schaffer, and Seabright (2006) have highlighted the ease with which a subjective ranking of constraints allows a comparison of the importance of different constraints, as in figure 1. This is not readily possible with objective indicators that measure
## Table 1. Most Severe Constraint Perceived by Entrepreneurs, by Region (percentage of firms)

<table>
<thead>
<tr>
<th>Region</th>
<th>Access to finance</th>
<th>Access to land</th>
<th>Corruption</th>
<th>Court system</th>
<th>Crime, theft, and disorder</th>
<th>Electricity</th>
<th>Inadequately educated workforce</th>
<th>Regulation</th>
<th>Political/macro-economic framework</th>
<th>Informal sector competitor practices</th>
<th>Tax administration</th>
<th>Tax rates</th>
<th>Telecommunications</th>
<th>Transportation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>15.2</td>
<td>4.3</td>
<td>5.8</td>
<td>0.8</td>
<td>7.6</td>
<td>17.2</td>
<td>4.4</td>
<td>7.8</td>
<td>7.6</td>
<td>9.2</td>
<td>3.4</td>
<td>9.8</td>
<td>1.1</td>
<td>5.7</td>
</tr>
<tr>
<td>East Asia and Pacific</td>
<td>22.8</td>
<td>4.6</td>
<td>4.1</td>
<td>0.3</td>
<td>3.2</td>
<td>10.2</td>
<td>8.0</td>
<td>8.9</td>
<td>4.7</td>
<td>17.3</td>
<td>2.3</td>
<td>8.5</td>
<td>0.0</td>
<td>5.1</td>
</tr>
<tr>
<td>Europe and Central Asia</td>
<td>15.9</td>
<td>2.6</td>
<td>7.0</td>
<td>2.0</td>
<td>2.5</td>
<td>5.5</td>
<td>9.7</td>
<td>9.3</td>
<td>10.7</td>
<td>11.1</td>
<td>4.5</td>
<td>17.6</td>
<td>0.0</td>
<td>1.7</td>
</tr>
<tr>
<td>L. America and Caribbean</td>
<td>8.5</td>
<td>1.3</td>
<td>7.9</td>
<td>1.7</td>
<td>7.4</td>
<td>6.2</td>
<td>6.7</td>
<td>11.1</td>
<td>18.9</td>
<td>13.0</td>
<td>4.7</td>
<td>10.2</td>
<td>0.0</td>
<td>2.4</td>
</tr>
<tr>
<td>South Asia</td>
<td>13.4</td>
<td>6.3</td>
<td>4.8</td>
<td>0.2</td>
<td>9.6</td>
<td>17.8</td>
<td>3.6</td>
<td>7.2</td>
<td>26.4</td>
<td>2.9</td>
<td>2.1</td>
<td>2.8</td>
<td>0.0</td>
<td>3.0</td>
</tr>
<tr>
<td>All</td>
<td>12.8</td>
<td>2.7</td>
<td>6.8</td>
<td>1.4</td>
<td>6.4</td>
<td>9.9</td>
<td>6.6</td>
<td>9.6</td>
<td>13.3</td>
<td>11.6</td>
<td>4.1</td>
<td>11.2</td>
<td>0.3</td>
<td>3.5</td>
</tr>
</tbody>
</table>

**Note:** Data from 41,207 firms in 91 countries for the years 2006-09.

**Source:** [http://www.enterprisesurveys.org](http://www.enterprisesurveys.org)
Table 2. Most Severe Constraint Perceived by Entrepreneurs, by industrial sector (percentage of firms)

<table>
<thead>
<tr>
<th>Sector</th>
<th>Access to finance</th>
<th>Access to land</th>
<th>Corruption</th>
<th>Court system</th>
<th>Crime, theft, and disorder</th>
<th>Electricity</th>
<th>Inadequately educated workforce</th>
<th>Regulation</th>
<th>Political/ macro-economic framework</th>
<th>Informal sector competitor practices</th>
<th>Tax administration</th>
<th>Tax rates</th>
<th>Telecommunications</th>
<th>Transportation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Textiles</td>
<td>14.6</td>
<td>1.3</td>
<td>5.7</td>
<td>1.2</td>
<td>5.3</td>
<td>6.9</td>
<td>7.1</td>
<td>11.2</td>
<td>13.9</td>
<td>15.3</td>
<td>3.9</td>
<td>11.2</td>
<td>0.1</td>
<td>2.2</td>
</tr>
<tr>
<td>Leather</td>
<td>14.3</td>
<td>9.5</td>
<td>4.8</td>
<td>0.0</td>
<td>0.0</td>
<td>14.3</td>
<td>23.8</td>
<td>14.3</td>
<td>0.0</td>
<td>4.8</td>
<td>4.8</td>
<td>10.5</td>
<td>0.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Garments</td>
<td>14.0</td>
<td>2.3</td>
<td>5.6</td>
<td>1.0</td>
<td>6.0</td>
<td>9.2</td>
<td>7.4</td>
<td>9.8</td>
<td>13.3</td>
<td>14.3</td>
<td>4.1</td>
<td>10.5</td>
<td>0.2</td>
<td>4.2</td>
</tr>
<tr>
<td>Food</td>
<td>12.2</td>
<td>2.2</td>
<td>5.9</td>
<td>1.5</td>
<td>6.9</td>
<td>12.0</td>
<td>5.8</td>
<td>9.6</td>
<td>11.9</td>
<td>12.7</td>
<td>4.1</td>
<td>10.8</td>
<td>0.2</td>
<td>4.2</td>
</tr>
<tr>
<td>Metals and machinery</td>
<td>13.6</td>
<td>2.7</td>
<td>7.7</td>
<td>1.7</td>
<td>5.4</td>
<td>8.6</td>
<td>8.9</td>
<td>10.0</td>
<td>12.8</td>
<td>8.6</td>
<td>4.4</td>
<td>13.0</td>
<td>0.0</td>
<td>2.6</td>
</tr>
<tr>
<td>Electronics</td>
<td>12.7</td>
<td>2.0</td>
<td>8.6</td>
<td>0.5</td>
<td>4.4</td>
<td>10.2</td>
<td>7.2</td>
<td>10.0</td>
<td>15.8</td>
<td>10.3</td>
<td>4.4</td>
<td>10.2</td>
<td>0.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Chemicals and pharma-ceuticals</td>
<td>10.9</td>
<td>1.7</td>
<td>8.4</td>
<td>1.8</td>
<td>4.8</td>
<td>6.5</td>
<td>5.4</td>
<td>13.2</td>
<td>17.4</td>
<td>12.5</td>
<td>4.4</td>
<td>9.8</td>
<td>0.0</td>
<td>3.2</td>
</tr>
<tr>
<td>Wood and furniture</td>
<td>18.1</td>
<td>2.9</td>
<td>3.1</td>
<td>0.5</td>
<td>1.8</td>
<td>9.3</td>
<td>11.5</td>
<td>10.4</td>
<td>4.8</td>
<td>13.0</td>
<td>4.6</td>
<td>16.3</td>
<td>0.0</td>
<td>3.7</td>
</tr>
<tr>
<td>Nonmetallic and plastic materials</td>
<td>16.6</td>
<td>3.9</td>
<td>6.9</td>
<td>0.9</td>
<td>4.4</td>
<td>10.4</td>
<td>5.8</td>
<td>8.5</td>
<td>13.5</td>
<td>11.9</td>
<td>3.4</td>
<td>9.3</td>
<td>0.1</td>
<td>4.4</td>
</tr>
<tr>
<td>Auto and auto components</td>
<td>11.5</td>
<td>3.6</td>
<td>1.8</td>
<td>0.6</td>
<td>1.2</td>
<td>2.4</td>
<td>10.3</td>
<td>12.7</td>
<td>2.4</td>
<td>7.3</td>
<td>9.1</td>
<td>35.8</td>
<td>0.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Other manufacturing</td>
<td>12.9</td>
<td>3.0</td>
<td>6.9</td>
<td>1.3</td>
<td>6.0</td>
<td>13.9</td>
<td>6.1</td>
<td>8.1</td>
<td>13.1</td>
<td>10.2</td>
<td>3.3</td>
<td>10.2</td>
<td>0.6</td>
<td>4.3</td>
</tr>
<tr>
<td>Retail and wholesale trade</td>
<td>12.0</td>
<td>3.2</td>
<td>7.4</td>
<td>1.5</td>
<td>8.0</td>
<td>8.0</td>
<td>5.7</td>
<td>9.4</td>
<td>12.4</td>
<td>11.8</td>
<td>4.5</td>
<td>12.1</td>
<td>0.5</td>
<td>3.5</td>
</tr>
<tr>
<td>Hotels and restaurants</td>
<td>13.9</td>
<td>3.8</td>
<td>5.8</td>
<td>0.9</td>
<td>6.5</td>
<td>16.3</td>
<td>8.1</td>
<td>6.3</td>
<td>10.2</td>
<td>9.1</td>
<td>4.4</td>
<td>12.3</td>
<td>0.0</td>
<td>2.5</td>
</tr>
<tr>
<td>Other services</td>
<td>11.6</td>
<td>2.6</td>
<td>6.3</td>
<td>1.2</td>
<td>6.3</td>
<td>10.1</td>
<td>7.3</td>
<td>9.6</td>
<td>15.1</td>
<td>10.4</td>
<td>4.4</td>
<td>11.5</td>
<td>0.6</td>
<td>3.1</td>
</tr>
<tr>
<td>Construction, transportation</td>
<td>13.0</td>
<td>3.1</td>
<td>8.8</td>
<td>1.6</td>
<td>6.6</td>
<td>6.0</td>
<td>7.4</td>
<td>9.9</td>
<td>14.7</td>
<td>10.5</td>
<td>4.1</td>
<td>10.1</td>
<td>0.3</td>
<td>3.9</td>
</tr>
<tr>
<td>All Sectors</td>
<td>12.8</td>
<td>2.7</td>
<td>6.8</td>
<td>1.4</td>
<td>6.4</td>
<td>9.9</td>
<td>6.6</td>
<td>9.6</td>
<td>13.3</td>
<td>11.6</td>
<td>4.1</td>
<td>11.2</td>
<td>0.3</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Note: Data from 41,207 firms in 91 countries for the years 2006-09.
various elements of the business climate in variable-specific units. For instance, it is much easier to ask directly firms to rank the perceived severity of the constraint posed by the power supply relative to corruption, rather than trying to rank it based on two objective measures such as the number of power outages relative to the amount of bribes paid. While over-optimism or pessimism may affect estimates of the absolute level of measured constraint severity, there is no reason to think that average differences between constraint rankings are likely to be biased. Thus subjective data may be helpful in shedding light on the relative importance of different constraints within economies. However, even if they can play an important complementary role, subjective indicators are probably less useful than objective ones in standard econometric analyses.

Recent Findings of the Enterprise-level Literature on Business Climate

This subsection summarizes the most important results of the recent business climate literature which relates firm performance to business climate indicators. Given that many studies have very specific and limited samples, one must be careful before drawing general conclusions. However, a large variety of samples can be shown to yield essentially similar or complementary results. The subsection is structured by type of constraints, looking in turn at infrastructure (electricity, telecommunications, transport, and water), competition and market regulation, financial constraints, and corruption and crime.

Infrastructure. A pioneering analysis of infrastructure indicators was done by Lee and Anas (1992) and Lee, Anas, and Oh (1996, 1999) for three developing
countries—Nigeria, Indonesia, and Thailand. Their analyses are not based on the standard enterprise survey data described above but on three dedicated surveys that were carried out in the late 1980s and early 1990s. The infrastructure information they collected, however, is very similar to the one available in the enterprise survey database. The results presented in these papers are dated by now but the concerns they address remain relevant today. There are large variations in the availability and quality of public infrastructure across countries, regions, and firm sizes. Lee, Anas, and Oh (1999) found that Nigeria tended to have a worse public infrastructure performance and a correspondingly higher incidence of private provision than Thailand and Indonesia, and speculated that the comparatively worse problems of Nigeria are related to the country’s then tighter restrictions on private provision arrangements. Aimed at protecting inefficient public suppliers, these restrictions prevented the emergence of private infrastructure provision regimes more efficient than the simple “one firm, one generator” model.

Small firms are disproportionately affected by infrastructure deficiencies. Lee and Anas (1992) find that, in the three countries they study, small firms depend more on public infrastructure and experience more power failures than larger firms because there are economies of scale in private provision of electricity and water: it is relatively cheaper for larger firms to provide their own power and water. This result finds support in Aterido, Hallward-Driemeier, and Pagès (2007) who use more than 80 enterprise surveys to examine deviations from the average ranking of perceived constraints. They find that small firms report electricity as a greater relative constraint than larger firms. Smaller firms are more likely to be in areas without access to electricity or to be dependent on an unreliable public grid, and lack the scale economies to operate a generator efficiently. Since a large share of new jobs is created by small firms, the negative impact of infrastructure deficiencies on employment creation is potentially huge. Regrettably, none of these papers attempts to measure the potential costs in terms of lost job opportunities of infrastructure constraints faced by small firms.

Infrastructure has a significant impact on enterprise productivity. The most severe constraint is electricity. Many developing countries are unable to provide their industrial sector with reliable electric power. Many enterprises in these countries have to contend with insufficient, poor-quality electricity and opt for self-generation even though it is widely considered a second-best solution. Table 3 shows the severity of electricity hazards across regions and per capita GDP levels as revealed by enterprise surveys for 104 countries in 2002–06. Survey data on the number of power outages are available for only 87 countries, on backup generators for 77 countries, and on cost of electricity for 34 countries (Alby, Dethier, and Straub 2010). As a result of the constraints faced by enterprises shown in table 1, many firms invest in backup power generators. On average, 31 percent of all firms own a generator (62 percent and 37 percent in South Asia and Africa,
respectively), with a large variance across firms in terms of number of power outages and generator ownership. These differences correlate significantly with firm size: 46 percent of large firms, 29 percent of medium-sized firms, and 17 percent of small-sized firms report owning a generator. The advantage of this study is that it develops a theoretical model, providing structure to estimate the impact of infrastructure deficiencies on firms’ input choices.

Dollar, Hallward-Driemeier, and Mengistae (2005), using survey data from Bangladesh, China, India, and Pakistan, find that, even after controlling for firm characteristics and region- or country-level effects, power losses have a significantly negative effect on total factor productivity. This seems to confirm the importance of electricity in poor countries and more generally the significance of infrastructure for explaining variation in productivity. Aterido and Hallward-Driemeier (2007) carry out a related study with particular focus on Africa. They are able to confirm that a higher incidence of power outages has a negative impact on employment growth. African firms seem to have adapted to this problem to some extent so that, for a given frequency of outages, employment growth in Africa is stronger than expected, relative to the rest of the world. This has partly to do with the comparatively high incidence of generator ownership in Africa, which reduces the impact of power shortages from the public grid (Foster and Steinbuks 2009). However, another reason seems to be that a higher frequency of outages seems to have contributed to a disproportional concentration of African employment growth in very small firms, which are less capital intensive and thus less vulnerable to power outages in terms of employment effects. Dollar, Hallward-Driemeier, and Mengistae (2005) draw on a sample of enterprise surveys from Bangladesh, Brazil, China, Honduras, India, Nicaragua, Pakistan,

Table 3. Access to Electricity by Firms across Regions and Country Income Groups

<table>
<thead>
<tr>
<th>Region</th>
<th>Percent of firms for which electricity is a major or severe constraint (%)</th>
<th>Average number of power outages</th>
<th>Percent of firms having more than 30 power outages (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe and Central Asia</td>
<td>8.5</td>
<td>9.72</td>
<td>5.7</td>
</tr>
<tr>
<td>Latin America</td>
<td>9.3</td>
<td>12.44</td>
<td>7.7</td>
</tr>
<tr>
<td>East Asia and Pacific</td>
<td>25.1</td>
<td>36.49</td>
<td>18.3</td>
</tr>
<tr>
<td>Mid. East and N. Africa</td>
<td>21.5</td>
<td>41.32</td>
<td>22.1</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>16.4</td>
<td>61.12</td>
<td>45.2</td>
</tr>
<tr>
<td>South Asia</td>
<td>43.0</td>
<td>131.74</td>
<td>49.0</td>
</tr>
</tbody>
</table>

**Country group by GDP per capita**

<table>
<thead>
<tr>
<th>GDP per capita</th>
<th>Percent of firms for which electricity is a major or severe constraint (%)</th>
<th>Average number of power outages</th>
<th>Percent of firms having more than 30 power outages (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>4.9</td>
<td>1.32</td>
<td>0.2</td>
</tr>
<tr>
<td>Upper-middle</td>
<td>8.3</td>
<td>13.02</td>
<td>6.2</td>
</tr>
<tr>
<td>Lower-middle</td>
<td>14.3</td>
<td>13.76</td>
<td>9.1</td>
</tr>
<tr>
<td>Low</td>
<td>26.4</td>
<td>64.08</td>
<td>34.1</td>
</tr>
</tbody>
</table>

and Peru to estimate the probability of exporting of a randomly chosen firm in a given city. They include “losses from power outages” as one business climate indicator and find that it has a negative and significant impact on the probability of exporting.

Infrastructure explains 9 percent of firm-level productivity, which is the second highest percentage after red tape, corruption, and crime in Escribano and Guasch (2005). In this careful econometric study using Guatemala, Honduras, and Nicaragua survey data, various productivity measures are regressed on infrastructure variables (average duration of power outages, number of days to clear customs for imports, shipment losses as fraction of sales, and a dummy for internet access) and controls. For the pooled sample, a 1 percent increase in the average duration of power outages decreases productivity between 0.02 and 0.10 percent, depending on the productivity measure used, and which mainly affects older plants. A 1 percent increase in the fraction of shipment losses will decrease productivity between 1.23 and 2.53 percent, most importantly in old and small firms. Firms with access to the internet are 11 and 15 percent more productive than those firms without access. Some of their results must be interpreted with caution. For instance, the huge impact of internet access on productivity suggests that this dummy functions as a proxy for better equipped, higher-technology firms rather than just representing internet access per se. This points to a limitation of the econometric methodology to address both production function inputs and the potential endogeneity of investment climate variables. Moreover, firm-level TFP measured as a residual is a questionable concept in the sense that it may relate to both positive and negative aspects, for example monopoly power, in which case results on the relationship between competition and productivity (see below) would be difficult to interpret.

The Escribano and Guasch (2005) methodology has been applied to many country data including Brazil, Chile, Costa Rica, Mexico, Turkey, and Southeast Asia. Escribano, Guasch, and Peña (2010) examine the influence of infrastructure on the average TFP of enterprises in 26 African countries and find that poor-quality electricity provision affects particularly poor countries but can also affect faster growing ones such as Botswana, Namibia, and Swaziland. Losses from transport interruptions affect mainly slow-growing countries, such as Madagascar or Kenya. Bastos and Nasir (2004) obtain similar results for transition countries (Moldova, Poland, Tajikistan, Uzbekistan, and Kyrgyz Republic). Infrastructure accounts for the second largest share of the variation in firm-level productivity, behind competition but before rent predation. However, their results are not robust because their two-step estimation is vulnerable to simultaneity bias and they do not control for country effects. While they may capture some genuine cross-country differences in all business climate categories, they are also vulnerable to bias if other cross-country effects (such as trade policy or...
political instability) influence productivity and are also correlated with their indicators.

A number of studies linking firm performance and infrastructure suffer from reverse causality. More infrastructure can cause the efficiency of firms to improve, but better economic performance may also attract more infrastructure. Datta (2008) exploits panel data from India to avoid this reverse causality problem. Investigating the effects of a highway improvement program on the productivity of firms, he argues that if “the precise route of the highway was not manipulated to include some intermediate areas (counties, districts, cities) and exclude others based on factors correlated with the outcomes of interest, the highway construction could be treated as exogenous to the areas that the highway runs through.” This allows for a difference-in-difference estimation strategy in which changes in relevant outcomes for affected firms are compared to the corresponding outcomes for firms whose location precluded their directly benefiting from the highway program. Since the highway improvement program in question uses the most direct routes between destinations, no opting-out was possible and no realignments was carried out; the areas in between destinations can indeed be viewed as a quasi-random selection of locations with existing highways to which the upgrade treatment was applied. Datta finds that enterprises that profited from the upgrade held significantly lower inventories, became less likely to report transport as a major or severe problem, and showed a greater propensity to change suppliers between the two years (suggesting that they found more suitable ones). He interprets this as evidence that improved highways facilitated productive choices, eased the extent to which infrastructure bottlenecks constrain firms, and allowed them to be more efficient.

Papers that find no significant effects of infrastructure on firm performance are in the minority and generally use specific samples or have methodological limitations. For instance, Commander and Svejnar (2007) use the BEEPS round II and III surveys to regress firm revenue on a number of controls and subjective business climate variables, including a composite infrastructure variable based on the questions from appendix B. They find that perceived infrastructure constraints have a negative and significant effect on firm revenue, but only without controlling for country fixed effects. They conclude quite generally that only country effects (due partly to differences in infrastructure, partly to other unobserved heterogeneities) have an impact, while within-country differences in infrastructure do not. This seems to be a premature conclusion given that significant within-country effects are significant in many other studies and that their sample is limited to Eastern Europe and Central Asia.

Competition and Regulation. The view that competition and entry should promote efficiency and prosperity has now become common wisdom worldwide (Aghion
and Griffith 2005). Generally speaking, we would expect a positive effect of competition on firm performance and a negative effect of excessive regulation. Studies based on business climate survey data have already confirmed this. They are mainly based on cross-country regressions. Enterprise survey panel datasets which could yield estimates of the impact of regulatory changes over time are lacking.

Using survey data from 60 countries, Carlin, Schaffer, and Seabright (2006) show that anti-competitive practices (as well as tax rates and tax administration, access to and cost of finance, and policy uncertainty and macroeconomic stability) are the most important business constraints in all countries. Gelb and others (2007) look more closely at tax administration and labor regulations and argue that policies become more serious determinants of the business climate at this stage, largely because the state has stronger capacity to implement them. Tax administration is primarily a problem in middle income countries, and the perception of labor regulations as a severe constraint increases with the GDP level of the country. Figure 2, based on 2006–09 enterprise survey data from 91 developing countries, shows the time spent by management dealing with regulators and tax inspectors, by firm size for each major geographical region.

**Figure 2.** Time spent with regulators, by firm size and geographical region

![Graph showing time spent with regulators by firm size and geographical region.](image)

*Note:* Data from 41,207 firms in 91 countries.
Competition has a significant positive impact on productivity (whether measured as the number of competitors in a main product line or as the importance of domestic or foreign competition to introduce new products or to reduce costs) and explains a far larger part of the variation in enterprise performance than rent predation or infrastructure. Many papers include measures of competition in their business climate regressions, including Bastos and Nasir (2004), Escribano and Guasch (2005), Beck, Demirgüç-Kunt, and Maksimovic (2005), Hallward-Driemeier, Wallsten, and Xu (2006) and Commander and Svejnar (2007).

Beck, Demirgüç-Kunt, and Maksimovic (2005) find that the “degree of legal obstacles” has a significant negative impact on enterprise productivity but regulation does not necessarily have a negative effect—and has a positive effect when it is consistently enforced. Aterido and Hallward-Driemeier (2007) and Aterido, Hallward-Driemeier, and Pagès (2007) find that consistent enforcement of regulations has a clear positive association with employment growth in most developing countries—though it is not significant for Africa—and is particularly marked for small firms. Both papers also obtain a generally positive effect of management time spent dealing with authorities, which they interpret as representing the benefit from obtaining public goods. On the other hand, pure red tape—for example unnecessary delays in customs—has a significant negative effect. A limitation common to these papers is the potential endogeneity of input choices that may bias the results directly (as for labor in Aterido, Hallward-Driemeier, and Pagès 2007) or indirectly if it affects productivity (as in Escribano and Guasch 2005).

**Financial Constraints.** The cost of finance and access to finance are often among the most severe constraints faced by enterprises. Across countries, the cost of finance is ranked above average in terms of severity in all country groups by Carlin, Schaffer, and Seabright (2006) and, on the African continent, it is the highest ranked constraint except in South Africa. The severity of the access to finance constraint declines with the GDP level of the country in Gelb and others (2007).

Within countries, enterprise size appears to be determinant as it influences the ability of obtaining credit from banks. Beck, Demirgüç-Kunt, and Maksimovic (2005) regress a firm level indicator of financial access on firm size using 54 datasets from the World Business Environment Survey (http://info.worldbank.org/governance/wbes/). Even after controlling for a country’s institutions, smaller firms report significantly higher financial obstacles than large firms. Ayyagari, Demirgüç-Kunt, and Maksimovic (2008) examine the financing constraints faced by enterprises using the same datasets and conclude that maintaining policy stability, keeping crime under control, and undertaking financial sector reforms to
relax financial constraints are the most effective ways to promote enterprise growth. Likewise, Aterido, Hallward-Driemeier, and Pagès (2007), on the basis of objective enterprise survey data, find that smaller firms have significantly less access to different forms of finance even when controlling for age, export status, ownership, and industry. The enterprise survey data indicates that small firms tend to finance a much smaller share of their investments with formal credits. Bigsten and Söderbom (2006) show that close to two-thirds of microfirms, but only 10 percent of large firms, are credit constrained in their sample of African countries. In their regressions, controlling for other important factors such as expected profitability and indebtedness, the likelihood of a successful loan application varies with firm size. World Bank (2008a) examines in more details the determinants and implications of lack of access to finance by enterprises. Also, within countries, access to finance is particularly problematic for less productive firms (Carlin, Schaffer, and Seabright 2006).

Financial obstacles have, in the vast majority of studies we reviewed, a negative significant effect on enterprise growth. Beck, Demirgüç-Kunt, and Maksimovic (2005) regress firm sales growth on subjective indicators of financial obstacles and controls. Six out of the eleven specific financial constraints indicators have a negative and significant impact—but their results could have an omitted variable bias. Moreover they do not calculate location-industry averages and their estimates are vulnerable to reverse causality at the firm level. Aterido, Hallward-Driemeier, and Pagès (2007) are more careful in reducing endogeneity at the firm level. They find that in general a higher share of investments financed externally is associated with greater employment growth. Aterido and Hallward-Driemeier (2007) find that a 10 percent increase in the share of investments financed through bank loans (equivalent to doubling the average share) is associated with a 3 percent increase in employment growth. This result is robust to alternative measures of finance, including formal bank financing of investment to trade credit among firms.

By contrast, Commander and Svejnar (2007) cannot find a significant effect of their subjective ‘cost of finance’ variable on firm revenue in their dataset from Eastern Europe and Central Asia. Dollar, Hallward-Driemeier, and Mengistae (2005) find no significant effect of the indicator ‘access to overdraft facility’ on productivity of firms in the garment industry, but in an expanded sample they do find a significant and strongly positive impact of the variable on annual sales growth. Dollar, Hallward-Driemeier, and Mengistae (2006) find a robust positive relationship between access to overdraft and the probability that a firm is an exporter. The study of the business climate in China by Hallward-Driemeier, Wallsten, and Xu (2006) yields no significant link between a variety of firm performance indicators and bank access which, as pointed out above, may largely be due to the peculiar nature of the Chinese state-owned banking sector which tends
to be relatively inefficient and subsidizes unsuccessful enterprises for political reasons.

**Corruption and Crime.** Crime, bribery, and corruption are identified as major problems for enterprises in less developed countries. Crime and corruption show up as important constraints in all country groups except the OECD, with crime a constraint in 25 percent of countries and corruption in 70 percent (Carlin, Schaffer, and Seabright 2006). Gelb and others (2007) shows that concern about corruption and crime tends to peak in the middle of the per capita income range. They interpret this as meaning that once economies overcome utmost poverty and the most basic limitations related to infrastructure, finance, and macroeconomic stability, problems of low administrative and bureaucratic capacity come to the forefront of firms’ concerns.

Recent studies that examine the relationship between firm performance and business climate indicators generally find significant effects for corruption and crime indicators. Fisman and Svensson (2007) use their Ugandan firm-level dataset for a study focused on corruption and its effect on growth. Their OLS and IV regressions of sales growth on a corruption indicator and a variety of controls show a “strong, robust, and negative relationship between bribery rates and the short-run growth rates of Ugandan firms, and...the effect is much larger than the retarding effect of taxation.” Keeping crime under control is one of the most effective ways to promote enterprise growth according to Ayyagari, Demirgüç-Kunt, and Maksimovic (2008). The enterprise productivity study by Escribano and Guasch (2005) includes “payments to deal with bureaucracy faster as percent of sales” and “number of criminal attempts suffered” as explanatory variables. The coefficient for the number of crimes suffered is significant and negative. However, the size of bribe payments has a robust positive relation with productivity. This may mean that firms that can afford to pay more bribes will tend to be more productive in the first place, reap productivity advantages from their payments, or both. But it certainly does not imply that the incidence of corruption should be seen as positive for productivity in general. Still, the difference in the sign of the corruption variable in the study is somewhat puzzling, and further research would be required to reveal the source of the difference (which could be genuine cross-country variation in the mechanisms of corruption, or related to more problematic endogeneity concerns).

Size of the enterprise matters. Aterido, Hallward-Driemeier, and Pagès (2007) find a significantly negative effect of a bribery dummy and other corruption indicators on the growth of small, medium, and large firms but a significantly positive effect for microfirms. This could mean that microenterprises find it easier to escape the attention of corrupt officials and therefore tend to grow faster than larger firms if the industry-location averages of corruption are higher. In their
study of the Chinese business climate, Hallward-Driemeier, Wallsten, and Xu (2006) find that objectively measured corruption matters a great deal for sales growth. Reducing the mean score of corruption by one standard deviation has a positive effect on sales growth by 6 percentage points. However, no significant effect of corruption can be shown for other firm-performance indicators such as productivity and employment growth. Regressing sales growth on corruption indicators and controls does not yield significance, though the coefficient has the expected negative sign in Beck, Demirgüç-Kunt, and Maksimovic (2005). The authors attribute this to multicollinearity in the sense that the “impact of corruption on firm growth is captured by the financial and legal obstacles.” Bastos and Nasir (2004) also find that their rent predation aggregate (which is meant to measure a combination of corruption and regulation) has a significantly negative effect, but it explains less variation of productivity than the infrastructure and competition measures.

Lessons and Ways Forward

The results of firm-level studies reviewed in the previous section relate enterprise performance to various business climate indicators, along with a series of controls for variables such as firm characteristics, industry, and country effects. These studies provide new evidence about one of the central assertions of the 2005 World Development Report on the business climate, namely that a good business climate drives growth by encouraging investment and higher productivity (World Bank 2004). At least four aspects of the business climate—infrastructure, finance, corruption and crime, and competition and regulation—have been shown to have a significant impact on firm performance.

The firm-level studies have already improved on the macroeconometric literature in a number of respects. They have shown that within-country heterogeneity is important. Variation in local business climate does indeed matter for explaining differences in firm performance. Much the same point is made by single-country, regional-business climate studies such as the China study by Hallward-Driemeier, Wallsten, and Xu (2006). Moreover the much larger sample sizes made possible by moving to a more disaggregated level allow for more robust results than in aggregate studies. The information obtained from the business climate surveys is also much more detailed and practical than aggregate indicators, allowing, for instance, insights about the variation of business climate effects across regions and different types of firms.

Building on this, it becomes possible to build a rich research program. In this section, we first outline the econometric issues and limitations of the current literature and derive the main lessons. We then highlight what, in our view, are the
most promising areas for future research. We conclude by opening the debate on potential improvements in the design of existing survey questionnaires.

**Econometric Lessons from the Literature**

The standard approach in the literature based on enterprise survey data has been to use regression analysis to identify which—if any—business climate indicators determine firm performance and to what extent. Almost universally, the basic specification of these regressions has been:

\[
\text{Firm Performance} = \beta_1 + \beta_2 (\text{Business Climate Indicators}) + \\
\beta_3 (\text{Firm Characteristics}) + \beta_4 (\text{Additional Controls}) + \epsilon
\]

When interpreting results from these regressions, it is important to keep some basic characteristics and limitations of the approach in mind. First, significant coefficients of the explanatory variables are only obtained if there is variation in these variables. The results obtained efficiently pinpoint existing bottlenecks explaining observed variations in firm performance, but they are less useful for identifying universal problems. For instance, Hallward-Driemeier, Wallsten and Xu (2006) find that access to banking services is not a significant determinant of firm performance in China. However, this does not mean that increasing the availability and efficiency of financial services is unimportant for improving Chinese productivity. As the authors point out, “it only means that the state-owned banking sector has not contributed significantly to regional firm growth.”

The fact that Chinese state-owned banking has not had a systematic impact on firm performance means that it does not show up as a determinant of actual variation therein. But the common lack of efficient banking services may still be responsible for suboptimal levels of firm performance throughout China. This methodological issue is particularly relevant for studies with small samples because expanding the number of observations will tend to introduce more variation and thus allow more general statements.

A related issue is that of “camels and hippos” raised by Hausmann and Velasco (2005) and discussed in Gelb and others (2007) and other papers. All results are necessarily based on the answers of existing firms that were interviewed. However, if one only interviews those present (“camels in the desert”), one may miss the crucial constraint (“water”) of those who have not entered (“no hippos in the desert”). A self-selected sample may imply a lack of variation in the explanatory variables that prevents us from noticing a critical constraint. Gelb and others (2007) argue that such self-selection is hardly ever complete (for example hippos can be expected to live in a water hole at the edge of the desert) and that firms that choose to enter in spite of serious constraints (which may force them into
costly evasive actions) will perceive them as particularly serious and thus introduce econometrically significant variation. Still, as it stands it is important to recognize that the econometric model above only informs us about the effect of constraints on the sample of existing firms. It is sometimes argued that the more interesting issue is rather the underlying industrial structure (for example the camel/hippo ratio in the desert) which should give away the most important constraint (that is the absence of hippos indicates that the main constraint is the lack of water). This, however, could only be addressed with completely different models such as “entry” models. Also one could think about exploiting a symmetrical issue, namely “exit.” Indeed specific type of firms may be more affected by changes in the environment (for example hippos in cases of severe drought). To our knowledge, this has not yet been addressed in the literature, probably partly because consistent panels are just becoming available, partly because of important attrition issues that need to be considered.

Another general methodological problem is that of multicolinearity. If regressors are correlated with each other, estimates will be inefficient and it may be impossible to know the importance of any one particular indicator since it may be serving as a proxy for other, more relevant variables. This is a particular problem with the business climate data, as many indicators are closely related. For instance, the prevalence of email usage may largely move with the quality of electricity supply. This counsels caution when interpreting very specific indicators, and emphasizes the importance of choosing a good regression specification. To some extent, variables such as “prevalence of email” should be seen as proxies for broader infrastructure factors. The solution chosen by Bastos and Nasir (2004) is to aggregate explicitly a number of specific indicators into broader measures (infrastructure, competition, etc.) in order to get clearer results at the loss of some (presumably misleading) detail. However, this makes the derivation of concrete policy implications more difficult.

Endogeneity—that is, a correlation between the explanatory variables and the error term—is more serious than multicolinearity because it causes not only inefficiency and interpretative difficulties, but bias and inconsistency of the estimates. The presence of endogeneity undermines the validity of estimated relationships between business climate indicators and firm performance.

It is unrealistic to assume that firm-level business-climate indicators are exogenous for a number of reasons. First, a major endogeneity problem arises if relevant explanatory variables are mistakenly omitted from the regression equation and also correlated with relevant included regressors. If this is the case, the estimated parameters of the included regressors will pick up some of the impact on the dependent variable of the omitted variables with which they are correlated. This will distort the estimates of the parameters of the included indicators.
Regressors, because they will now capture both their own effect and part of that of the correlated omitted variables.

Second, better subjective and objective business climate indicators may be associated with better performing firms, not because they cause such firms to be more productive, but on the contrary because “an inherently more efficient firm can work within the exogenously given environment to reduce inspections, power losses or days for customs clearance or phone lines” (Dollar, Hallward-Driemeier, and Mengistae 2005). Similarly not only may better suited business environments cause firms to be more efficient, but inherently more efficient firms may also be more likely to have the necessary resources to identify and (re)locate to better suited environments. At the aggregate level, inherently more prosperous regions may have greater political clout to obtain infrastructure and other business climate improvements from government. If one cannot fully control for these reverse causality factors, estimates of the effect of the business climate on firm performance will be biased.

The firm-level business-climate literature suggests various measures to limit the endogeneity bias:

- Regressions on single business climate indicators are likely to produce biased and inconsistent parameter estimates due to omitted variables. A sufficiently broad array of indicators and controls should therefore be used in regression equations. The selection of regressors should go from general to specific (Carlin, Schaffer, and Seabright 2006).
- Objective indicators are generally preferable to subjective ones as explanatory variables because they are less vulnerable to measurement error and reverse causality (Bertrand and Mullainathan 2001).
- Using location-industry or industry averages instead of (or as instruments for) the firm-level objective indicators can help alleviate endogeneity due to reverse causality. See for instance Escribano and Guasch (2005); Hallward-Driemeier, Wallsten, and Xu (2006); Dollar, Hallward-Driemeier, and Mengistae (2006); Commander and Svejnar (2007). All use location-industry averages that exclude the respective firm. The idea is that while better region-industry business-climate indicators should explain variation in firm performance, individual firm performance has virtually no impact on the average-indicator. This alleviates direct reverse causality.
- Country-level effects should be controlled for—either with country dummies or with specific country-effects variables—to avoid a contamination of the business climate coefficients with correlated but unobserved country-level effects on firm performance.
- In the absence of panel data, an approach similar to Miguel, Gertler, and Levine (2005) might be useful to alleviate some endogeneity problems. They
try to explain industrialization—measured as growth in manufacturing employment—at the district level in Indonesia over a 10-year period with social capital indicators at the beginning of the period, but find no effect. A similar approach could be taken with indicators from enterprise surveys.

- A simple two-step estimation procedure estimating TFP as the residual of a production function, then attempting to explain TFP with business climate indicators, is potentially vulnerable to simultaneity bias. The problem is that, in most cases, the production function inputs will be correlated with the indicators because the business climate influences not only productivity per se, but also input choices of firms. This means that in the production function regression, the error term (that is TFP) is likely to be correlated with the regressors (labor and capital) via the business climate, leading to bias. This approach should thus be avoided. Escribano and Guasch (2005) have suggested alternative procedures.

- Considering moreover that TFP, initially defined as a “measure of our ignorance” in aggregated data, is a problematic concept when applied at the firm level, it might be safer though to concentrate on simpler outcomes (for example employment or sales in levels or growth rates) and even to start by deriving firms’ input choices from underlying structural models (see Alby, Dethier, and Straub, 2010).

On this last point, a related issue arises with the quality and relevance of the performance proxies used as dependent variables (productivity, profit, sales growth, etc.). Going into the details of the literature on this topic would take us beyond the scope of this paper, but we should note that measures of firm-level productivity are much more likely to run into problems and generate biases since the very construction process of these variables make them likely to be correlated with policy shocks and managerial decisions (see Katayama, Lu, and Tybout 2009). This is not to say that alternative proxies (for example profit, sales, or employment growth) are completely free of problems (see Del Mel, McKenzie, and Woodruff 2007) but again, in many cases, they appear to be preferable.

**A Possible Research Agenda**

There remain a number of areas in which additional research could bring interesting results. At the theoretical level, we need to develop a better understanding of the link between firm choices and the business climate in developing countries. A growing body of empirical research is relating cross-country differences in economic outcomes, such as productivity or output per capita, to differences in policies and institutions that shape the business environment. Some empirical research
has also shed light on the determinants of productivity at the firm level and the evolution of the distribution of productivity across firms within each industry—modeling decisions about investment, R&D, employment, and so on, which hinge on the type of constraints revealed by the existing surveys (things like credit constraints, infrastructure bottlenecks, level of competition in goods and labor markets, volatility of macroeconomic conditions, entry costs, commitment and enforcement problems or information issues).

Resource misallocation can lower aggregate TFP. Hsieh and Klenow (2009) and Bartelsman, Haltiwanger, and Scarpetta (2009) investigate the effect of firm-level policy distortions on aggregate outcomes. Hsieh and Klenow use microdata on manufacturing establishments to quantify the potential extent of misallocation in China and India versus the United States. They measure sizable gaps in marginal products of labor and capital across plants within narrowly defined industries in China and India compared with the United States. When capital and labor are hypothetically reallocated to equalize marginal products to the extent observed in the United States, they calculate manufacturing TFP gains of 30–50 percent in China and 40–60 percent in India. Bartelsman, Haltiwanger, and Scarpetta (2009) use harmonized data from 24 OECD and Eastern European EU member countries on firm-level variation within industries. They find substantial variation in the within-industry covariance between size and productivity across countries, but this covariance varies significantly across countries and is affected by the presence of idiosyncratic (that is firm-level) distortions.

For developing countries, the literature on firm choices of formality found for example in Rauch (1991), Straub (2005), or De Paula and Sheinkman (2008) can be useful in this context. Together with tools from industrial organization and contract theory, this approach provides a good basis for formalizing insights on market behavior in developing countries. Additionally, results could then be used to understand the very different shapes of firms’ distributions we see across countries, for example in terms of size, productivity, or exporting behavior, and guide the empirical applications.

At the empirical level, some of the most interesting insights in the firm-level literature on developing countries have come from studies examining the interactions of business climate indicators with firm characteristics or with each other. For instance, Aterido and Hallward-Driemeier (2007) interact business climate measures with firm sizes to obtain more detailed results on the impact of the business environment on the performance of different types of firms. Honorati and Mengistae (2007), examining the interplay of regulation, infrastructure, financial constraints, and corruption, obtain interesting results, for instance that all three aspects have significant influence on Indian industrial growth, yet their effect depends on the incidence of corruption. Most existing firm-level studies
have not considered these types of interactions, and more work in this direction could deliver interesting findings laying the groundwork for more precisely targeted policy recommendations.

A key research goal highlighted by a number of authors is that when more survey rounds become available, proper panel data regressions could test for the impact of changes in the business climate on productivity, factor returns, and growth. For instance, whereas current microeconomic studies predominantly aim to assess the variation in firm performance due to local and cross-country variations in existing constraints, panel data could allow an assessment of the impact of changes (reforms) in the shape of different constraints on firm performance, as well as on the entry and exit patterns of firms. However, with only three survey rounds available at most, it is still relatively early for these types of studies.

Even the standard methodological approaches have not yet made full use of the large enterprise survey database. For instance, no best-practice study (properly accounting for endogeneity) of the relationship between firm productivity and the business climate has been carried out for the full, up-to-date enterprise survey database. This has only been done with employment growth as dependent variable (Aterido, Hallward-Driemeier, and Pagés 2007). A very comprehensive study of infrastructure in Africa has finally been published (Africa Infrastructure Country Diagnostic, forthcoming) but not enough analysis is available on the impact of infrastructure on manufacturing firm productivity on this continent (Bigsten and Söderbom 2006). On the other hand, there is also scope to carry out detailed country studies such as that of China by Hallward-Driemeier, Wallsten, and Xu (2006) or those of India by Honorati and Mengistae (2007) and Amin (2007). It is generally easier to interpret correctly econometric results in single-country studies because outcomes are easier to connect to real-life circumstances and complementary data.

Durlauf, Kourtellos, and Tan (2008) argued that the effect of institutions is likely to be through their influence on proximate growth determinants (factor accumulation, in this case) rather than through their effects on technological innovation. It would be interesting to explore this question further with enterprise data. So far, only a few papers have used measures of capital (or human capital) accumulation as a dependent variable and there has been no systematic comparison to the results for TFP.

Future studies should make sure to test extensively the robustness of their results and if possible improve on the methodology in a more fundamental way. This is because even best-practice precautions against endogeneity—such as using location-industry averages as instruments for firm-level indicators, regressing on multiple business climate indicators at a time, and controlling for the current country, region, and industry effects—leave regressions vulnerable to inconsistency and bias, as several researchers have pointed out. For instance,
location-industry averages are used as instruments to alleviate endogeneity stemming from reverse causality. Yet such endogeneity effects can persist at a more aggregate level as well because of endogenous placement decisions of firms and policy endogeneity. For instance, industry-location averages may yield a strong relationship between firm performance and average quality of telecom services in a specific industry and region. But as Carlin, Schaffer, and Seabright (2006) point out, regions that are prosperous for a variety of other reasons for which it is not realistically possible to control econometrically also happen to have higher levels of telecom services. To counter this effect, some studies (for example Hallward-Driemeier, Wallsten, and Xu 2006) include additional city information and sector dummies to help control for those more macroissues that affect both the business climate and the firm. But the inability to control sufficiently for all factors implies that the endogeneity problem is likely to persist to some extent. In light of this, it is clear that the need arises both for more modeling efforts (for example Alby, Dethier, and Straub 2010) and for more creative instrumental strategies. Some examples of the latter can be found in the literature—for example in Duflo and Pande (2007), who use geographical data to instrument for the endogenous placement of infrastructure, or Datta (2008) and Gibson and Rozelle (2003), who take advantage of the seemingly exogenous placement of road works in specific contexts to assess their impact.

One final comment on translating the findings of this literature into policy. Regression results based on data provided by firm managers are often straightforwardly translated into policy advice, for example to increase competition and lower regulation. The underlying assumption is that changes in the business climate which improve firm performance will translate into broad social benefits. Yet this may not always be the case. In order to reduce the risk of any negative impact on social welfare, it is wise to consider possible competing interests when examining policy implications. For instance, regulations may impact on firm productivity negatively but provide benefits to nonmanagerial social groups.

**Improving Questionnaire Design**

At a fundamental level, it may also be worthwhile to rethink the standard enterprise survey questionnaire which determines the raw data on which all analyses are based. For instance, in the era of cellphones—which are particularly important in many developing countries—the focus on mainline telephone services is anachronistic and misleading. With regard to infrastructure indicators, Straub (2008) makes a number of suggestions for more detailed questions such as firms’ access to alternative transport modes (railways, airports, roads, etc.) or the ownership of vehicles.
There appear to be many holes in the information provided. In electricity, for example, most information is on quality (outages and cost thereof) but basic information on cost and availability of service, such as average cost of a kilowatt-hour of electricity from the public grid or cost of power generators, is absent. Similarly, for water, information is sought on the sources of provision but it should be complemented with the respective average unit costs. In transport, data on the possibility to access different types of services (roads, railways, etc.) together with an assessment of their unit cost and quality, and the ownership of different types of vehicles, would make it possible to assess the significance of the transport mix chosen by firms. In the case of telecommunications, mobile telephony is completely absent from existing surveys. Here again, data on access, unit cost, type (such as gathering information on markets, money services, etc.), and quality of services derived from mobile phones would be necessary. One also wonders why questions geared at the use of the internet are restricted to the subsample of service firms.

Finally, a few key dimensions would need to be added. First, information on the institutional nature of service providers and regulatory arrangements would be crucial from a policy perspective. Moreover in a context where the geographical dimension is increasingly recognized to be important (Gibson and McKenzie 2007; Straub 2008), data need to be spatially referenced. Obviously the practical task of gathering this type of data (including in particular several hours spent with directors and managers of firms, who often have imperfect knowledge about the things they are asked to report) implies a trade-off between being exhaustive and collecting data of good quality.
### Appendix A: Studies of the Relationship between Enterprise Performance and Business Climate

<table>
<thead>
<tr>
<th>Paper/Dataset</th>
<th>Type of analysis</th>
<th>Dependent variables</th>
<th>Independent variables</th>
<th>Results/critique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bastos and Nasir (2004)</td>
<td>Estimation of TFP and subsequent regression of TFP on three measures (rent predation, infrastructure, and competition)</td>
<td>Ln (TFP)</td>
<td>Rent predation variable, based on:</td>
<td>Results</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The business climate variables have the expected signs and are jointly significant.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Using Kruskal methodology, competition is found to explain far more variation in firm-level productivity than infrastructure, which in turn explains more variation than rent predation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Infrastructure variable, based on:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Country dummies are not included (see critique of Commander and Svejnar 2007).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Two-step estimation procedure vulnerable to simultaneous equation bias as outlined by Escribano and Guasch (2005).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Competition variable, based on subjective estimates of importance of domestic/foreign competition to introduce new products/reduce costs.</td>
</tr>
<tr>
<td>Using BEEPS II extended dataset for Kyrgyz Republic, Moldova, Poland, Tajikistan &amp; Uzbekistan</td>
<td>Based on:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Firm age, Exports (as percent of sales), Foreign ownership.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To find out whether firm size determines perceptions on financing, legal, and corruption constraints to doing business, the authors carry out OLS regressions of firm-level financing and legal and corruption indicators on firm size, controlling for country level financing, legal, and corruption constraints.

To see whether variations in perceived obstacles can explain firm sales growth, the authors carry out OLS regressions of firm growth on business climate indicators (one at a time), controlling with industry dummies, firm characteristics, and country random effects.

To explore how the effects of business climate indicators differ by firm size, size

32 subjective business climate variables in financial, legal, and corruption categories. Firm growth (percent change in firm sales over past three years).

Business climate variables (1 to 4 scale):
• summary financing obstacle
• summary legal obstacle
• summary corruption obstacle.

29 subjective and three objective IC indicators for more specific constraints within the three summary categories.

Controls: Ownership (government and foreign); Exporter status; number of competitors; industry dummies; country dummies (inflation; GDP; GDP per capita; GDP growth).

Results
1. Firm perception of the financing and corruption obstacles it faces relates to size, with smaller firms reporting significantly higher obstacles than large firms. Smaller firms report lower legal obstacles than do larger firms, but these differences are not significant.
2. When entered individually, all three summary obstacles have a negative and significant effect on firm growth. Entered jointly, financing and legal obstacles are both significant and negative, but corruption loses its significance. Entering the 32 individual obstacles one at a time, some of the financing and corruption variables, but none of the legal ones are significant.
3. The authors find evidence that financial
controls are added as explanatory variables.

obstacles have a much greater impact on the operation and growth of small firms than on that of large firms. Critique Study uses few objective business climate indicators (3 out of 35). Results on the 32 specific business climate variables are based on regressions of firm growth on each of the variables in turn, therefore likely omitted variable bias.


GLS and Levinsohn/Petrin production function estimation of TFP in the garments industries of all countries. TFP is then regressed on the logs of a set of business climate variables and controls. Regression of factor rewards in garments industries on the same variables plus firm characteristics. The For pooled data (all industries):
• Sales growth
• Annual growth rate of fixed assets
• Annual growth rate of employment.

For garments industry only:
• TFP
• average wage
• average profit.

Business climate variables:
• log(custom days export)
• log(custom days import)
• log (power loss)
• log (days to get phone)
• log (overdraft facility).

Instrumented by city-sector averages.

Control variables (not all in every regression): log of: distance from market; distance from port;

<table>
<thead>
<tr>
<th>Paper/Dataset</th>
<th>Type of analysis</th>
<th>Dependent variables</th>
<th>Independent variables</th>
<th>Results/critique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dollar, Hallward-Driemeier, and Mengistae (2005) “Business climate and Firm Performance in Developing Economies” Using four World Bank Enterprise Survey datasets for Bangladesh, China, India, and Pakistan</td>
<td>GLS and Levinsohn/Petrin production function estimation of TFP in the garments industries of all countries. TFP is then regressed on the logs of a set of business climate variables and controls. Regression of factor rewards in garments industries on the same variables plus firm characteristics. The For pooled data (all industries):</td>
<td>• Sales growth</td>
<td>• log(custom days export)</td>
<td>obstacles have a much greater impact on the operation and growth of small firms than on that of large firms. Critique Study uses few objective business climate indicators (3 out of 35). Results on the 32 specific business climate variables are based on regressions of firm growth on each of the variables in turn, therefore likely omitted variable bias.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Annual growth rate of fixed assets</td>
<td>• log(custom days import)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Annual growth rate of employment.</td>
<td>• log (power loss)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• log (days to get phone)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• log (overdraft facility).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Instrumented by city-sector averages.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Control variables (not all in every regression): log of:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>distance from market;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>distance from port;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Results robust to inclusion</td>
</tr>
</tbody>
</table>
hypothesis is that factor rewards will be higher when business climate is better.

Regression of sales growth, growth in fixed assets, and growth in employment in all industries (pooled dataset) on business climate variables and controls.

All regressions carried out for the full sample and a subsample of small firms.

Escribano and Guasch (2005)
"Assessing the Impact of the Business climate on Productivity Using Firm Level Data: Methodology and the Cases of Guatemala, Honduras and Nicaragua"
Using ICA datasets for Guatemala, Honduras, and Nicaragua

Paper objective is to develop a consistent econometric methodology to be used as a benchmark for evaluating the impact of IC variables on productivity at the firm level. The methodology is then applied in the cases of Guatemala, Honduras, and Nicaragua.

The econometric analysis consists of a variety of regressions of productivity measures on business performance.

Productivity at enterprise level (10 different measures)

Business climate variables: instrumented by region-industry averages.
Red tape, corruption, and crime:
- no. of days spent in inspection and regulation related work
- fraction of sales undeclared to tax authorities
- payments to deal with bureaucracy faster (% of sales)

Results
Analyses which use two-step procedure (first estimate firm productivity, then regress this measure on IC variables) are likely to suffer from simultaneity bias. Therefore paper proposes three different methods to estimate productivity.

Use of four categories of business climate variables deemed important for

of country dummies which shows that business climate at local level is important.

Critique
Regressions with TFP potentially vulnerable to simultaneity bias.
However, results largely confirmed with alternative firm-performance dependent variable.

Continued
climate indicators and a set of controls. Results are also analyzed by country, age, and size of firms.

† no. of criminal attempts suffered infrastructure
† average duration of power outages (log)
† days to clear customs for imports (log)
† shipment losses (% of sales)
† dummy for internet access.

Quality, innovation, and labor skills:
• fraction of computer controlled machinery
• fraction of total staff engaged in R & D
• dummy for ISO quality certification
• fraction of total staff with secondary or higher education
• dummy for training beyond “on the job” training.

Finance and corporate governance:
Guatemala, Honduras, and Nicaragua: (a) red tape, corruption, and crime; (b) infrastructure; (c) quality, innovation, and labor skills; and (d) finance and corporate governance.

Estimates show consistently high impact of business climate (especially red tape, corruption and crime, and infrastructure) on productivity. Overall, it accounts for over 30 percent of productivity.

Critique It is not clear that endogeneity of production function inputs is adequately solved. Moreover firm-level TFP is a fuzzy concept that may cover many aspects, both positive and negative.

Appendix A. Continued

<table>
<thead>
<tr>
<th>Paper/Dataset</th>
<th>Type of analysis</th>
<th>Dependent variables</th>
<th>Independent variables</th>
<th>Results/critique</th>
</tr>
</thead>
</table>


Regression analysis to test for a significant impact of corruption and taxes on firm sales growth, controlling for other factors

Sales growth: \[
\log(\text{sales}_{1997}) - \log(\text{sales}_{1995}) / 2
\]

Business climate variables:
- reported bribe as share of sales
- reported tax as share of sales

Instrumented by location-industry average.
Index (0–5) of availability of public services (electricity, water, telephone, waste disposal, paved roads).
Index of regulation (log of 1 + percentage of senior management time spent dealing with regulation).

Control variables:
- Ownership (foreign > 50%);
- log of firm age;
- (log of) sales in 1995;
- trade (firm exports and/or imports).

Results Both taxation and bribes are found to have a robust, significantly negative impact on short-run sales growth; the retarding effect of bribes is thereby stronger than that of taxes. Foreign ownership has a positive impact on sales growth, as does trade status at least in one specification.

Continued
### Appendix A. Continued

<table>
<thead>
<tr>
<th>Paper/Dataset</th>
<th>Type of analysis</th>
<th>Dependent variables</th>
<th>Independent variables</th>
<th>Results/critique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dollar, Hallward-Driemeier, and Mengistae (2006) “Business Climate and International Integration” Using eight World Bank Enterprise Survey datasets for Bangladesh, Brazil, China, Honduras, India, Nicaragua, Pakistan, and Peru</td>
<td>Probit regression of export status on business climate indicators and control variables (geographic, sector, firm characteristics). The aim is to relate business climate to the probability that a randomly chosen firm in a particular city exports. Country dummies are used in some specifications to obtain analysis of within-country variation.</td>
<td>Indicator variable of whether firm exports or not</td>
<td>Business climate variables: Three objective business climate indicators: days to clear customs; access to overdraft; losses from power outages. One subjective business climate indicator: whether managers thought government services inefficient. The authors use location averages to instrument the variables.</td>
<td>Results Paper finds that a sound business climate—as reflected in low customs clearance time, reliable infrastructure, and good financial services—makes it more likely that domestic firms will export, enabling the more productive firms to expand their scale and scope. The empirical link is largely robust to the inclusion of country dummies (at least joint significance and some individual significance remains) showing that local factors matter for the business climate.</td>
</tr>
<tr>
<td>Carlin, Schaffer, and Seabright (2006) “Where are the Real Bottlenecks? A Lagrangian Approach to Identifying Constraints on Growth from Subjective Survey Data” Using 60 World Bank</td>
<td>Overview of descriptive statistics of subjective business climate indicators. Development of model of the firm to predict relationship between reported TFP (defined for manufacturing firms using TFP residuals or the firms’ self-reported technological level)</td>
<td></td>
<td>Business climate variables: 17 subjective indicators of the severity of different business climate constraints on a four/five point scale (see appendix B in the text). Control variables: Country dummies; ownership</td>
<td>Results Descriptive statistics show that physical infrastructure rarely rates highly as a constraint; problems with licensing and customs affect relatively few countries (especially CIS); crime and/or corruption show up as</td>
</tr>
</tbody>
</table>

---

The World Bank Research Observer: vol. 26, no. 2 (August 2011)
Enterprise Survey datasets from BEEPS 2002, 2004 and 2005, and PICS 2000–05. Constraints and the characteristics of firms. Regression of TFP on business climate and controls. Important constraints in all groups of countries except the OECD: seven dimensions of the business environment that are ranked as of greater than average importance in all country groups: anti-competitive practices, tax rates and tax administration, access to and cost of finance, and policy uncertainty and macroeconomic stability. Regressions yield results in line with model predictions. Between-country regressions show negative and significant effects of telecom, electricity, transport, customs regulation, Mafia, land title, and land access indicators. Once country effects are controlled for, however, customs regulations, transport, and legal system indicators have perverse positive signs. Authors
argue that this is due to endogeneity bias. Finance has significant negative impact, as predicted because it does not have public good characteristics but instead inherently unproductive firms are rationally denied credit (and complain about this).

Critique
First, instead of objective indicators, the authors use subjective ones which are particularly vulnerable to the endogeneity effects they allege. Second, regressions only use one business climate indicator at a time, exposing them to omitted variable bias. Third, the posited relationship between firm performance and perceived indicator severity can only be shown for customs regulation and finance but not for any of the other disaggregated indicators. When avoiding
these problems, other authors do find nonperverse, significant effects even when employing country dummies. Fourth, authors do not use industry-location averages of their regressors, which could lessen firm-level endogeneity biases.


Regression of four different firm performance variables on largely objective business climate indicators (measured as city-industry averages) and controls

- Sales growth
- Investment rate
- Productivity
- Employment growth.

Business climate variables:
- Mean loss of sales due to transport/power outages
- Mean share of labor that uses computers
- Mean share of R&D staff in labor
- Mean regulatory burden
- Mean corruption
- Mean share of nonpermanent labor
- Mean bank access.

Control variables:
Ownership (domestic private/foreign); logs of firm age + 1 and firm age + 1 squared; log lagged sales; log lagged

Results
- Ownership significant, foreign ownership more so than private domestic
- No evidence that physical infrastructure matters significantly, but technological infrastructure does (expected given that road and power are good quality in China)
- Labor market flexibility weakly significant
- No evidence that average access to finance in a region and industry affects performance

Continued
As expected, ownership has strong effects on firm performance. Relative to state ownership, domestic private ownership is associated with a higher sales growth rate and investment rate. Effect of foreign ownership even larger. There is no evidence that physical infrastructure affects firm performance but the impact of technological infrastructure appears to matter significantly. Labor market flexibility matters weakly.

Critique The two-step TFP regression is vulnerable to simultaneity bias (see Escribano and Guasch 2005). However, the

<table>
<thead>
<tr>
<th>Paper/Dataset</th>
<th>Type of analysis</th>
<th>Dependent variables</th>
<th>Independent variables</th>
<th>Results/critique</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>employment; log city population and GDP per capita; city and industry dummies.</td>
<td>(expected given inefficient Chinese bank sector) • government regulatory burden and corruption strongly significant.</td>
<td></td>
</tr>
</tbody>
</table>
A descriptive overview of firm-level employment growth and business climate data from over 100 countries, focusing in particular on differences by firm size.

Regression of employment growth on IC constraints controlling for a variety of firm characteristics (especially size) and other factors.

- Firm employment growth.
- Business climate variables: Country-city-sector-size averages of the following variables:
  - Finance:
    - firm has overdraft facility
    - percent of sales sold on credit
    - percent of working capital financed externally
    - percent of investments financed externally.
  - Regulation:
    - log of days to get an operating license in last two years
    - percent of management’s time dealing with government regulation
    - log of days spent on inspections last year
    - log of average days to obtain imports last year

Results
Significant differences across size categories of firms—both in terms of objective conditions faced by firms and nonlinearities in the impact of these conditions. Low access to finance, corruption, poorly developed business regulations, and infrastructure bottlenecks shift downward the size distribution of employment. Low access to finance and ineffective business regulations reduce the growth of all firms, especially micro- and small firms. Corruption and poor infrastructure create growth bottlenecks for medium and large firms.

The results also reinforce...
## Appendix A. Continued

<table>
<thead>
<tr>
<th>Paper/Dataset</th>
<th>Type of analysis</th>
<th>Dependent variables</th>
<th>Independent variables</th>
<th>Results/critique</th>
</tr>
</thead>
</table>

- log of average days to get exports through customs last year
- log of total days spent on labor inspections last year.

**Corruption:**
- firms in comparable activities bribe to get things done (yes/no)
- bribes as percent of sales to get things done by similar firms
- similar firms give gifts to officials (yes/no)
- percent of government contracts on bribes by comparable firms

**Infrastructure:**
- power outages during the last year (log days)
- percent of sales lost due to power outages last year
- log of days with no water last year.

The importance of differentiating the impact across size classes of firms that allow for the micro firms (less than 10 employees) to be different from small firms.

A weak business climate reduces overall employment in the business sector. Firms may be confined to industries with limited innovation and growth opportunities. In addition, a larger share of firms may remain informal or semi-informal, reducing the capacity of the state for collecting taxes and paying for fundamental inputs for development such as education.

Descriptive overview of how African employment growth rate compares to the rest of the world, and how specific business climate constraints differ across regions and types of firms.

Regression of employment growth and other outcome variables on a set of business climate indicators and controls for two country samples (African countries and other developing countries).

Results
Firms in Africa face greater obstacles in terms of finance, infrastructure, public services, and governance. The more challenging business environment conditions do not translate into lower average growth compared to other developing countries, but they are associated with shifting down the firm size distribution, lowering the relative growth of larger firms, or in some cases expanding microfirms. This may be because there are incentives to remain small, e.g. because bad

Employment growth
• capital intensity
• change in capital intensity.

Business climate variables:
• share of investments financed with bank loans
• days without power
• management time with officials
• frequency of bribes to “get things done.”

Alternative specifications include: days to clear import customs; consistency of enforcement of regulations; share of sales on credit.

Control variables (not all in every regression): firm size, firm age, export status, foreign ownership, sector controls; survey dummies, country controls.

Continued
transport infrastructure creates demand-pockets for small suppliers, being small and informal minimizes contact with corrupt state.

**Results**

Country effects matter for firm performance but differences in the business environment constraints observed across firms within countries do not. Foreign ownership found to have positive effect on firm performance, but domestic private ownership not. Export orientation found to have positive effect only in simple specification, not if authors control for ownership.

**Critique**

Paper uses only subjective, perception based business climate indicators. Objective business climate indicators, such as the time required to clear customs, have been found to be significant even
• 10 heritage foundation indices of economic freedom.

Control variables (not all in every regression): levels of capital and labor inputs; categories of ownership (privatized, new private, foreign); export orientation of firm; log of exports/sales. Note that main controls are replaced by instrumental variables.

with the inclusion of country dummies (e.g. by Dollar, Hallward-Driemeier, and Mengistae 2005). Subjective business climate indicators may be suboptimal because systematic variations in perceptions in the cross-country dataset are largely a function of broad business confidence related to macrofactors such as political and macroeconomic stability or the financial system. Such country level effects are largely captured by country dummies. Once firm characteristics (such as size or age) are also controlled for, the remaining variation in perception based indicators of specific business climate constraints may be largely due to random factors such as managers’ personality. This could explain why these subjective indicators do not
Honorati and Mengistae (2007)
“Corruption, the Business Environment, and Small Business Growth in India”
Using data from Indian Firm Analysis and Competitiveness Surveys for 2002 and 2005

Descriptive analysis of objective and subjective data. Regressions analysis to examine the effects of corruption, labor regulation, access to finance, and the quality of power supply on the growth of manufacturing businesses in India.

Annual sales growth

Business climate variables:
- lagged profitability (finance proxy)
- lagged indebtedness (finance proxy)
- indicators of corruption, labor regulation, and power shortages.

Controls:
industry, state, and year dummies; initial size.

Results
The authors find a pattern whereby the better performing states are also better in every important aspect of their business environment: low-income, low-growth states have the worst indicators of all institutional variables except for labor regulation. Regressions show that average business growth rate is lower where labor regulation is greater, power shortages are more severe, and financial constraints stronger. Influence of each of these three factors on business growth depends on the incidence of corruption, sales growth is constrained by cash-flow only in businesses that are not affected by labor regulation.
power shortages, or corruption. The authors interpret this as indication that corruption is a proxy for something more fundamental than the payments of bribes, namely the quality of property right institutions. Their results are consistent with the view that the quality of property rights institutions exerts more abiding influence on economic outcomes than the quality of contracting institutions.
Appendix B: Infrastructure Variables — Subjective and Objective Questions

In the core survey questionnaire, there is one subjective perception variable for three types of infrastructure: electricity, transport, and telecommunications:

- Rate whether the following issues are a problem for the operation and growth of your business on a five point scale from ‘No obstacle’ up to ‘Very severe obstacle’: (a) Telecommunications, (b) Electricity, (c) Transportation.
  [14 other non-infrastructure issues are also listed, including customs/trade regulation, labor regulation, etc.]

There are also a number of objective indicators:

- During how many days last year did your establishment experience the following service interruptions, how long did they last, and what percent of your total sales value was lost last year due to: (a) power outages or surges from the public grid? (b) insufficient water supply? (c) unavailable mainline telephone service?
- Does your establishment own or share a generator? If yes, what percentage of your electricity comes from your own or a shared generator?
- What share of your firm’s water supply do you get from public sources?
- What percentage of the value of your average cargo consignment is lost while in transit due to breakage, theft, or spoilage?
- Does your enterprise regularly use email or a website in its interactions with clients and suppliers?
- Based on the experience of your establishment over the last two years, what is the actual delay experienced (from the day you applied to the day you received the service or approval) and was a gift or informal payment asked for or expected to obtain for each of the following? (a) A mainline telephone connection, (b) An electrical connection, (c) A water connection, (d) . . . [three other non-infrastructure issues].

Some changes to the core instrument have been made over time. For infrastructure, for instance, there were two additional questions which, unfortunately, were omitted after 2006:

- What is your average cost of a kilowatt-hour (kWh) of electricity from the public grid?
- If yes [on generator ownership], what was the generator’s original cost to your establishment?
These questions were omitted even though, in 2006, on average, close to half of all the firms surveyed experienced more than 30 power outages per year and 31 percent of firms owned a backup power generator. Specific national surveys may add infrastructure questions to augment the core survey.

Notes

J.-J. Dethier (corresponding author) is a research manager and Max Hirn is a junior professional officer, both at the World Bank. Stéphane Straub is professor at the Toulouse School of Economics, Arqade. Their email addresses are jdethier@worldbank.org; mhirn@worldbank.org; and stephane.straub@univ-tlse1.fr. The authors are grateful to Howard Pack, Emmanuel Jimenez, and three anonymous reviewers for helpful comments, and to Rabia Ali for outstanding research assistance.

1. Typically GDP per capita (for example in Acemoglu, Johnson, and Robinson 2001), GDP per worker (for example in Hall and Jones 1999), or their growth rates (for example in Knack and Keefer 1995 or Mauro 1995).

2. As discussed in Pritchett (2000), it is unlikely that infrastructure investment results from the equalizing of costs and benefits, as governments are usually not profit maximizers. The exact mapping between investment and the value of infrastructure created rather involves issues such as lack of efficiency in public investment, corruption, and wasteful public spending.

3. Not to be confused with the World Business Environment Survey (WBES) mentioned above, which was a one-off project in 1999–2000.

4. In both types of surveys, key descriptive statistics are calculated (number of firms that own a generator or a private well; production time, sales value lost due to public infrastructure interruptions etc., or both). Stratification and sample sizes are also similar in both cases (a few hundred enterprises per country and year, stratified by industry, region or city, and firm size).

5. World Bank (2004), in figure 6.4, reports that a greater percentage of large firms rank infrastructure constraints as major or severe. This is not inconsistent with the fact that smaller firms perceive infrastructure to be a greater relative constraint. There may be structural reasons—such as larger firms’ greater demands on various business climate dimensions—leading larger firms, on average, to report higher absolute constraint rankings in the various categories.

6. Transport is considered a severe constraint in only a handful of poor or war-torn economies, as well as in Ireland. Telecommunications is not perceived as a significant constraint, possibly indicating the extent to which the rapid spread of cellphones has resolved most communications problems. This underlines the need to update the survey questionnaire which until now refers to “mainline” telephone services only.

7. The other indicators may also capture some additional variation from unobserved variables. In essence, this means that there may still be some omitted variable bias that distorts the estimated parameters, or alternatively there may be no bias but the included variable may only be a proxy for the actual cause of the productivity effect.

8. The marginal impact of power outage duration on productivity could be much higher after a certain threshold, which may not be reached in middle-income countries.

9. Africa (10), South Asia (5), East Asia (7), Latin America and the Caribbean (7), OECD Europe (6), Central and Eastern Europe (8), South Eastern Europe (8), and the Commonwealth of Independent States (11).

10. South Africa is an exception: the constraints ranked most highly there are labor regulation, skill shortages, macroeconomic stability, and crime.

11. ‘How problematic is financing for the operation and growth of your business: (1) no obstacle, (2) a minor obstacle, (3) a moderate obstacle or (4) a major obstacle?’

12. ‘(i) Are collateral requirements of banks/financial institutions no obstacle, a minor, a moderate or a major obstacle?: (ii) Is bank paperwork/bureaucracy no obstacle?: (iii) Are high interest
rates no obstacle?; (iv) Is the need of special connections with banks/financial institutions no
obstacle?; (v) Is banks’ lack of money to lend no obstacle?; (vi) Is the access to foreign banks no
obstacle?; (vii) Is the access to nonbank equity/investors/partners no obstacle?; (viii) Is the access to
specialized export finance no obstacle?; (ix) Is the access to ease finance for equipment no obstacle?
(x) Is inadequate credit/financial information no obstacle?; (xi) Is the access to long term finance no
obstacle?’

13. Except when the variable is entered separately from the other business climate variables, which renders it vulnerable to omitted variable bias.

14. Their variable is the city-industry share of the corruption score, which is constructed as the
principal component of two variables: the ratio of bribes to sales plus the share of contract value
used as a bribe to get a business contract.

References


Alby, P., J.J. Dethier, and S. Straub. 2010. “Firms Operating under Infrastructure and Credit Constraints
in Developing Countries. The Case of Power Generators.” Paper presented at the International
econ.worldbank.org/conferences/ icied).

4314, World Bank, Washington.

177–200.

Infrastructure on Employment Growth: Does Sub-Saharan Africa Mirror Other Low-Income

Growth: The Impact of Access to Finance, Corruption and Regulations Across Firms.” Institute for
the Study of Labor (IZA) Discussion Paper 3138, Bonn.

Ayyagari, Meghana, Asli Demirgüç-Kunt, and Vojislav Maksimovic. 2008. “How Important are
Financing Constraints? The Role of Finance in the Business Environment.” World Bank Economic

A. Philippe, and S. Durlauf, eds., Handbook of Economic Growth. Amsterdam: North Holland:
473–552.

Differences in Firm Dynamics.” In T. Dunne, J.B. Jensen, and M. Roberts, eds., Producer Dynamics:


Development and Cultural Change 52:159–85.
41:6–10.
Private Alternatives and Policy Options.” Discussion Paper INU 98. Infrastructure and Urban
Development Department, World Bank, Washington.
Lee, K.S., A. Anas, and G. Oh. 1996. “Cost of Infrastructure Deficiencies in Manufacturing in
Washington.
———. 1999. “Costs of Infrastructure Deficiencies for Manufacturing in Nigerian, Indonesian and
93(7–8): 950–64.
928. Yale University Economic Growth Center.
They Matter and How Do Firms Respond to Them?” Institute for the Study of Labor (IZA)
Discussion Paper 1424, Bonn.
33–47.
Papers, 10(1): 40–70.


In the past decade, the World Bank has promoted improving business environments as a key strategy for development, which has led to a significant effort in collecting surveys of the investment climate at the firm level across countries. The author examines the lessons that have emerged from the papers using these new data. The key finding is that the effects of business environments are heterogeneous and depend crucially on industry, initial conditions, and complementary institutions. Some elements of the business environment, such as labor flexibility, low entry and exit barriers, and a reasonable protection from the “grabbing hands” of the government, seem to matter a great deal for most economies. Other elements, such as infrastructure and contracting institutions (that is, courts and access to finance), hinge on their initial status and the size of the market. JEL codes: K2, L5, L1, O1

Recently policymakers and multinational organizations have increasingly focused on a sound investment environment as a strategy for economic development (Stern 2002; World Bank 2005). It is difficult to define the “investment climate” or the “business environment” precisely. But Stern (2002) notes that it is the “policy, institutional, and behavioral environment, both present and expected, that influences the returns, and risks, associated with investment” in a specific location. In other words, the business environment covers whatever external environment that affects the returns and risks faced by investors. This general definition includes three broad categories. The first category covers macroeconomic aspects such as fiscal, monetary, and exchange rate policies, which clearly affect investors’ returns. High tax rates, for example, would lower return, while
inflation would increase the variability of returns. The second category includes governance, institutions, and political stability. Rule of law, for instance, affects investors’ decisions about how much to invest and what organizational form it should take. Institutions also include informal ones, such as the general level of trust, social capital, and social network (North 1990; Knack and Keefer 1997; Zak and Knack 2001; Shirley 2008), which would facilitate new transaction relationships and, therefore, firm expansion. The final category includes infrastructure necessary for productive investment, such as transportation, electricity, and communications. I discuss the effects of the second and third categories of the business environment, which include government policies and behavior related to the provision of infrastructure, tax burdens, protection of property rights, labor, and entry regulations (World Bank 2005).

Identifying the effects of the business environment is difficult. The first difficulty concerns measurement problems (for example, eliciting truthful responses about corruption). For a long time the available data have been only at the country level. But multicollinearity about various aspects of the business environment is particularly severe at this level. For instance, among the commonly used cross-country International Country Risk Guide (ICRG) governance indicators, the correlation coefficient between the rule of law and the control of corruption is 0.62 between 1996 and 2007. Identifying the business environment effects at the cross-country level thus becomes particularly difficult. This shortcoming can be partially overcome with firm-level data, which often allow us to go further than cross-country and cross-industry data. Of great help are the vast variations within a country. The average tax burden in inland Chinese provinces, for instance, is twice that of coastal areas (Cai, Fang, and Xu forthcoming). To understand the effects of tax burden on firm performance using firm-level data, one can take advantage of within-country variations, while holding the legal system (and therefore de jure institutions) constant. Moreover, some key measures of a country can be obtained only with firm-level data (Bigsten and Soderbom 2006). For example researchers have often used dispersion of productivity within an industry of a country to capture industry-level competition, which can be constructed only by exploiting firm-level data. And to understand what types of firms are credit constrained or particularly vulnerable to government expropriation, only firm-level data can be relied on. Unfortunately very little comparable firm-level data was available in developing countries—until now.

The past 10 years have witnessed an explosion of firm-level evidence on the effects of business environments, an explosion due partly to the data efforts of the World Bank, mainly the World Bank Enterprise Surveys (also known as the Investment Climate Surveys). These data efforts lead to consistent measures of business environments, which, in turn, usher in a substantial new literature on
how business environments affect development at the firm level. To push the research forward, it is helpful to review what lessons have been learned.

I will survey the evidence on how business environments affect economic development. Given the vastness of this task, one has to be selective. Thus I will focus on firm-level research related to the World Bank Investment Climate Surveys—with occasional discussions of outside firm-level research that are particularly relevant to our topic. This would automatically leave out studies based on the complementary Doing-Business Survey of the World Bank. A benefit of this omission is that we can focus exclusively on de facto institutions or regulations (in the case of Investment Climate Surveys) while ignoring new issues related to de jure institutions or regulations (in the case of Doing-Business Surveys) (Hallward-Dreimier, Khun-Jush, and Pritchett 2010). Several topics are too broad to cover, but there are excellent surveys already done and not much is lost in skipping them. The skipped topics include the effects of large-scale privatization and the effects of reforms in access to finance and corporate governance.6 Almost all investment climate surveys are cross-sectional in nature, with the notable exception of the Business Environment and Enterprise Performance Survey (BEEPS) in Eastern Europe and investment climate surveys in a few countries such as India. However, since institutions and investment climates change slowly, the attempt to identify the effects of the investment climate relying on within-country across-time variations in a short timespan is unlikely to go far (Griliches and Hausman 1986). Not surprisingly, I am unaware of any solid firm-level empirical studies using panel investment climate surveys that have been conducted so far.

The most common mode of identification with cross-sectional firm-level data is to construct proxies of the local or national business environment by using city- or country-level measures of access to finance, tax burden, corruption, labor flexibility, and so on, and then relate them to firm performance. In the case of country-level measures of business environments, multiple-country data have to be used. A variety of robustness checks are often used to ensure the robustness of the key findings. Often, omitted city-level or country-level proxies (such as the level of development and other aggregate-level variables) are added to ensure the robustness of the results (Cai, Fang, and Xu forthcoming). Alternatively instruments for key business environment measures are used. An example is to use the distance from surrounding enforcement offices to instrument the enforcement of regulations (Almeida and Carneiro 2009). Another common method of identification in this literature is the difference-in-difference approach using cross-country cross-industry data, in which the more disaggregated outcomes are regressed toward country dummies, industry dummies, and an interaction term of country-level treatment with an industry-level sensitivity variable.7 The idea is that “more sensitive” industries should be more affected by the treatment. By controlling for country and industry dummies, all country- and industry-specific
factors are controlled for, which makes this method more convincing and less subject to the omitted variable bias than most other cross-sectional estimations. If the interaction term proves to be significant, more sensitive industries are indeed responding as predicted to the treatment, supporting the notion of causality from the treatment to the outcome, especially for the sensitive type. Often sensitivity checks are applied—in which more interaction terms of the treatment with other industry-level variables are used—to make sure that it is indeed “the sensitivity variables” rather than the robust-check variables that cause the original interaction term to be statistically significant.

Given the cross-sectional nature of the investigations, and questionable validity of many exclusion restrictions needed to construct instrumental variables, most research using the investment climate data cannot establish causality convincingly. Invariably the estimations suffer from reverse causality, omitted variable bias, and other issues. The results should therefore be interpreted as a collection of correlations. To the extent that the results are robust across similar contexts, or consistent with plausible theories, the conclusions are more credible.

Since the conclusions related to the effects of the business environment for each individual study have to be tentative, it is even more imperative to summarize existing evidence and seek a coherent storyline to tie them together. As the reader will see, the body of correlations gathered from the studies does point to a plausible story: the effects of the business environment vary across industry, complementary institutions, and initial conditions. Some elements of the business environment turn out to loom large in most economies, such as a basic protection of property rights against government expropriation, labor flexibility, and low entry and exit barriers, which are found invariably to be important in explaining economic performance in various economic contexts. Other elements, such as infrastructure and contracting institutions, hinge critically on initial conditions. Infrastructure, for instance, is found to matter much more in countries with a low initial stock of infrastructure, while the quality of courts and access to finance are more important in richer countries.

I will first offer a simple framework of the effects of business environment reforms. I then summarize the evidence from specific areas of reforms: physical infrastructure, property rights, labor regulation, then entry regulation. In the final section I conclude and offer policy implications.

A Simple Framework of Heterogeneous Effects of Business Environment Reforms

The effects of the business environment often differ by specific contexts and there are often country-specific bottlenecks (Kremer 1993; Shleifer 2005). For my purpose,
it is useful to consider three types of fundamentally different government–business relationships and associated business environments to attain desirable social outcomes: market discipline, private litigation, and public enforcement through regulation (Shleifer 2005). This order represents an increase in public control over economic activity and increasingly state-dominated business environments, highlighting a delicate tradeoff between disorder and dictatorship. The further down the list, the lower is the chance of disorder and the stronger the danger of dictatorship. The optimal strategy will depend on specific economic and institutional circumstances and may differ even across industries within the same country.

**Market Discipline**

Market discipline—relying mainly on market competition between firms without depending on either litigation or government regulation—is the best strategy when it proves to be sufficient to control disorder and to avoid Hobbesian anarchy, such as through the reputation mechanism and the natural death of inefficient firms. The lack of need for regulating entry for most industries is a case in point. Most entering firms are small, and they cannot survive when they prove to be inefficient in satisfying customer needs with low costs. So we can rely on market competition (without explicit government regulations) in shaping entrepreneurial qualities.

**Private Litigation**

Another option for enforcing good conduct is through the legal system, such as using litigation and courts. Courts have the advantage of potentially apolitical or experienced judges in dealing with specialized economic cases. But courts entail disadvantages as well. The judges can be subverted through bribes, can be influenced by politics when they are appointed by the government, and the strong and not the just may win the cases because of unequal distributions of resources. Moreover, in many developing countries, formal and lengthy procedures hamper the effectiveness of dispute resolutions. This legal solution would likely fail when there is a severe unequal distribution of resources and when historical heritages and the level of development do not allow a well-functioning legal system.

**Public Enforcement**

Public regulation could partially solve this problem. The advantages are that regulators, being experts, can impose better rules and that the government can provide incentives to ensure socially desirable outcomes. The disadvantages are two. Industries featuring concentrated interests (relative to consumers) can
capture regulators to preserve monopoly power and prevent entry (Stigler 1971; Peltzman 1976). Regulators can also abuse power to pursue self-interests. Since the check on government and regulators are particularly weaker in developing countries, regulations there are less attractive in general.

Although Shleifer’s framework offers useful guidelines for the optimal choice of business environments, there is no one-size-fits-all recipe for the choice (Kremer 1993; Rodrik 2007; Lin 2009). Countries often differ in areas with the largest reform payoffs, and there are often development “bottlenecks,” which conjures up the image of the famous failure of the space shuttle Challenger: with thousands of components, it “exploded because it was launched at a temperature that caused one of those components, the O-rings, to malfunction” (Kremer 1993). When production technologies feature strong complementarities for production factors or elements of business environments, a bottleneck such as poor infrastructure would reduce the return to all other production factors. This, in turn, would lead to lower incentives for workers to improve skills and to invest in human capital, which further reduce the output of the local economy and make infrastructure investment increasingly nonlucrative. Thus a bottleneck may trap a local economy in a poor equilibrium.9

A manifestation of the bottleneck hypothesis is the existence of policy complementarity (Kremer 1993). Do we have any supporting evidence in the development context? A piece of such evidence is the complementarity between entry regulation and labor flexibility in India. The effects of India dismantling the License Raj (that is, a system of central controls regulating entry and production activity in manufacturing) in the 1980s and 1990s were found to depend on a state’s labor regulations. After painstakingly careful empirical work relying on distinct timings of the national reform in different states, Aghion and others (2008) concluded that License Raj had no effects on average. However, after delicensing, industries located in states with pro-employer labor market institutions grew significantly faster than those in pro-worker environments. There was, therefore, complementarity between product market deregulation and labor market flexibility.

Infrastructure

Policymakers and development economists often view infrastructure (for example road, power, communication, and customs) as necessary for economic development. Good infrastructure allows firms to have lower transport and communication costs and therefore lower total costs to compete with their rivals and to export. A larger extent of the market due to a better infrastructure also facilitates a greater scope for specialization, which further reduces unit production costs. Increasing complexity of transactions enabled by better infrastructure facilitates
the need for new institutions (for example the courts), which then reduce future agency costs for trade (Demsetz 1967). Yet how infrastructure affects firm performance remains understudied, though several recent studies do shed light on this key issue.

Infrastructure is found to be the most important factor in explaining firm performance in Bangladesh, China, Ethiopia, and Pakistan (Dollar, Hallward-Driemeier, and Mengistae 2005). Using the World Bank Enterprise Survey data for these countries, these authors study how the business environment affects firm performance, including total factor productivity (TFP), wages, profits, growth rates of output, employment, and fixed assets. The measures of business environments include several proxies for infrastructure (for example custom efficiency, power loss, and the number of days to install phones), the share of firms with overdraft access, and the number of times per year that they are visited by government inspectors. The business environment affects TFP because better local governance allows the same bundle of inputs to produce more outputs due to lower transportation and transaction costs and a better protection of property rights.

Why does a better business environment lead to higher wages and profits? For countries in the same specialization cone, factor endowment decides trade structure (Schott 2003). To compete with more efficient producers (with better business environments), countries with worse business environments can only afford lower factor prices to stay competitive—resulting in lower wages (for labor) and profits (for capital owners).

Why do better business environments lead to higher growth rates? A better business environment causes higher returns to capital, which, in turn, engenders a higher investment and growth rate initially before eventually reaching a steady state. Since firm-level measures of the investment climate may be endogenous—for example more profitable firms may have a better connection with government officials and may therefore face systematically more or less government harassment—Dollar, Hallward-Driemeier, and Mengistae use the location-sector average of the investment climate proxies to measure the local business environment. They also control for country dummies. This approach mitigates two issues: (i) firm measures of business environment may be closely related to omitted variables at the firm level; (ii) firm answers on the investment climate may be responses to rather than causes of firm outcomes. However, the issue of omitted location-sector-level variables remains—the local investment climate proxies may merely represent other local omitted factors. With this caution in mind, the authors find broad patterns that business environments improve productivity and give workers and investors higher returns and higher growth rates. But among the business environment indicators, the three indicators of infrastructure (power loss, phone days, and custom delays) are the most important. Similarly infrastructure (as captured by power) enhances firm performance in Bangladesh (Fernandes 2008).
Good infrastructure is also found to facilitate international integration. Using a sample of firms in Bangladesh, Brazil, China, Honduras, India, Nicaragua, Pakistan, and Peru, Dollar, Hallward-Driemeier, and Mengistae (2006) examine how the business environment affects international integration. Relying on within-country variations by controlling for country effects—and using the city-level average of the investment climate measures as proxies of the local business environment—the authors find that good infrastructure explains foreign ownership and exporting.

The effects of physical infrastructure seem to differ by countries. China is richer than most of the countries mentioned earlier that feature positive infrastructure effects, and China has invested a large amount of money on physical infrastructure. Using within-China variations, Hallward-Driemeier, Wallsten, and Xu (2006) find that their proxies of physical infrastructure at the city level are not significantly associated with firm performance. Thus the positive association between infrastructure and firm performance seems to be particularly strong in countries with a worse stock of infrastructure—due perhaps to its decreasing marginal return.

Property Rights

Cross-country evidence suggests that countries with worse property rights tend to have lower aggregate investments, worse access to finance, and slower economic growth (North 1990; Knack and Keefer 1995; La Porta and others 1997, 1998, 2000; Acemoglu, Johnson, and Robinson 2001). However, micro support for the positive relationship between property rights and economic performance would be helpful in strengthening the case. With only macro evidence, it could be institutions causing growth or institutions following growth. Indeed, Glaeser and others (2004) find that human capital is a more basic source of growth than are institutions, and that poor countries get out of poverty through good policies, often initiated by dictators, and subsequently improve their political institutions.11 Further complicating the matter, property rights at the country level are often subjective and not precisely defined and measured. The ICRG measure of corruption, for instance, is ordinal, going from 1 to 6, and the increases of one point at different initial points do not have the same meaning. How these numbers are constructed remains a black box. Yet micro measures of corruption—such as the ratio of bribes to sales—are directly comparable across countries. Indeed, whether tax or corruption are more damaging, an issue of great concern for the literature of corruption, is more naturally studied in a micro setup, in which corruption can be measured as bribes over sales, with the same unit as tax burdens (that is, total taxes paid over sales). Their coefficients are then directly comparable.
Recent firm-level studies of property rights in transition suggest that the importance of property rights likely depend on the stage of transition. In the early stages, property rights are found to be overwhelmingly important by Johnson, McMillan, and Woodruff (2002), who use a 1997 sample of firms from Poland, Romania, Russia, Slovakia, and Ukraine to study the relative importance of property rights and finance. The sample consists of relatively small yet profitable firms (with 7 to 270 employees). After-tax profits over sales, ranging from 5.7 percent in Slovakia to 21 percent in Russia, are high because of unfulfilled market niche, entry barriers, and immature transitions. A nice feature of this paper is their objective measures of property rights at the firm level, which rely on whether firms pay extra-legal payments for licenses, for protection, and for services, and whether courts can or cannot enforce contracts. Access to finance is measured as whether firms had collateral and had loans in the previous year. At the country level, there is a strong association between property rights and reinvestment rate (that is, the share of firms’ profits that are used for investment). Russia and Ukraine fare the worst, while Poland and Romania fare the best in measures of property rights. Indeed the reinvestment rate in the latter group is on average one-third higher. Further regression analyses of firm reinvestment rates on firm-level perception of property rights robustly show that reinvestment rates are significantly related to the perception of secure property rights, but not with firm access to finance. Since the perception of property rights and the access to finance are all measured at the firm level, there might be firm-level characteristics that are omitted that cause artificial correlation of the key variables and the outcome variable. The authors deal with this by demonstrating the robustness of their key results, which survive with controls for industry characteristics, country characteristics, and manager traits. A caveat is that uncontrolled firm characteristics and reverse causality can also explain the pattern.

Are property rights always more important than finance? Not necessarily. McMillan and Woodruff (2002) conjecture that the relative importance of market-supporting institutions (that is, the courts and finance) should rise over the stages of economic development. As an economy gets richer, transactions become more complicated, specialization goes further, and it becomes more difficult to confine related producers in a single location and to rely only on personal relationships to sustain their business dealings. A reliable court system would thus be needed to enforce contracts based on such arms-length relationships. Moreover, since richer countries feature a larger scale of production, retained earnings would not suffice for further expansion, and external finance becomes necessary.

Examining the importance of property rights and finance at a more mature stage of transition, Cull and Xu (2005) confirm this conjecture with a large
sample of Chinese firms in 2000–02. The property rights measures consist of two components: the government expropriation measures (that is, the percent of sales spent on informal payment to government officials and the likelihood of government officials helping instead of hindering firms) and measures of ease of contract enforcement (that is, the percent of firms’ disputes resolved through the courts, the dummy of a firm signing formal contracts with clients, and the likelihood of a court upholding a firm’s legal rights in commercial disputes). The finance measures include access to bank finance, trade credit, and the collateral requirements (as a share of loan amounts). Lacking a big bang, the uncertainty on property rights was less severe in China than in eastern European countries. With the transition into a mature stage, China features strong competition and low profit margins, making external finance important. Indeed, Cull and Xu find that external finance was relatively more important in China than in the early stage of transition in Russia and the eastern European transitional countries; the external finance variables are found to be statistically significant and to have explanatory power in China, in comparison to their lack of explanatory power in the post-communist countries in eastern Europe. However, property rights remain important, especially for smaller firms. The empirical method of Cull and Xu (2005) are very similar to McMillan and Woodruff (2002), partly in order to be comparable. The caveats about omitted variables and reverse causality thus apply to this study as well.\textsuperscript{12}

The Mechanisms of Property Rights Effects

Several recent firm-level studies shed light on the mechanisms through which property rights affect firm performance. First, in a sample of 30 industrialized and developing countries, the protection of property rights, as measured at the country level, is associated with a better availability of external finance (Demirgüç-Kunt and Maksimovic 1998).\textsuperscript{13} When property rights are well protected, information disclosure about firm performance and fund uses are more adequate, so banks are therefore more likely to make loans, shareholders more willing to invest, and abuse of company funds more likely to be detected and punished. Similarly, when courts function well, creditors are more likely to get their loans back when firms underperform and declare bankruptcies, and are therefore more willing to lend.

A better protection of property rights also leads to a better asset allocation. Claessens and Laeven (2003), using cross-country firm-level data, find that industrial sectors that use relatively more intangible assets (as proxied by the importance of intangible assets for the corresponding industry in the United States) develop faster in countries with a better protection of property rights.\textsuperscript{14} They also confirm that the protection of property rights boosts the return to investments in

Colin Xu
intangible assets. Their results are more convincing than many existing studies in that they control for both country and industry fixed effects, therefore holding constant all factors that are country- and industry-specific. Why does property rights protection lead to a better asset allocation? Without a proper protection of intellectual property rights, firms fear expropriation of investment in intellectual property and intangible assets. They then invest more in tangible assets, which are less subject to government expropriations since the resale value upon exit would be higher for tangible than for intangible assets. The extra reliance on property rights protection for intangible-asset-intensive industries explains why the ratio of intangible assets is more associated with firm growth in countries with a better protection of intellectual property rights.

A better protection of property rights is also associated with a higher share of firms registered as limited liability corporations, which is an organizational innovation to reduce the risks faced by investors and to increase the ease with which to obtain external finance. What business environments lead to a higher tendency of incorporation? Demirgüç-Kunt, Love, and Maksimovic (2006), using a sample of firms in 52 countries from the World Business Environment Survey, find that businesses are more likely to choose the corporate form in countries with more developed financial sectors, better legal systems, stronger investor rights, lower regulatory burdens and corporate taxes, and efficient bankruptcy processes. In addition, the marginal return of incorporation is higher in countries with good financial and legal institutions—corporate firms grow faster in such environments.

Are All Property Rights Created Equal?

Acemouglu and Johnson (2005) unbundle institutions into “property rights institutions” and “contracting institutions.” Property rights institutions capture how much private property is secure from the “grabbing hand” of the state, for example through outright expropriation or bribe extraction. Contracting institutions capture the effectiveness of institutions that are used to resolve disputes between private contracting parties, such as the courts and the judicial system. Based on cross-country evidence, they find that property rights institutions tend to be far more important than contracting institutions in facilitating economic development. Only property rights institutions are consistently statistically significant and quantitatively important, whereas contracting institutions are not. Their interpretation is that it is easier for private parties to use alternative mechanisms to get around the contracting issues, but it is harder to avoid government expropriations.

The unbundling conjecture and the relative importance of property rights over contracting institutions are confirmed by Beck, Demirgüç-Kunt, and Maksimovic (2005), a study that relies on the World Bank Environment Survey data between 1996 and 1999 from 54 countries. In their main specification, they relate firm
growth rates to standard firm- and country-level controls, along with ordinal measures of firm-perceived degree of obstacles in financing, corruption, and the legal system. To avoid collinearity, they enter each firm-perceived obstacle one at a time. Their regression results show that while proxies of property rights institutions (for example general bribes, bribes to bank officials, and managerial burdens in dealing with regulators) are negatively associated with firm growth, the speed of the court (that is, a proxy of contracting institutions) in resolving disputes is not significantly so. To allow firm-perceived obstacles to be endogenous—firm-level determinants of growth may directly determine firms’ perceptions of the degree of an obstacle in a particular aspect—they instrument firm-perceived obstacle with country-level measure of institutions and find the results to be robust. This instrument variable approach hinges on the assumption that country-level measure of institution affects firm growth only through the corresponding firm-perceived obstacle.

In contrast, findings from a large firm-level dataset in China around 2003 suggest that both property rights institutions and contracting institutions are important (Cull and Xu 2005)—both sets of institutions are statistically and economically significant in predicting firm reinvestment rates.15

The Effects of Corruption on Firm Performance

A visible symptom of bad property rights institutions is corruption—the abuse of public office for personal gains. How does corruption affect firm performance? On the one hand, the "grease view" implies that corruption acts like an auction that allocates licenses and government contracts to the highest bidders, which is efficient since the most efficient firms can afford the highest bribes (Myrdal 1968; Lui 1985). On the other hand, corruption is potentially much more dangerous (Shleifer and Vishny 1993). It diverts resources from public uses such as the provision of public goods. Moreover, when regulators are decentralized and uncoordinated, the cumulative bribe burden on private agents may become excessive, and efficient economic activities such as foreign direct investment entry may not occur, resulting in efficiency losses. In addition, by its nature, corruption entails secrecy. To avoid exposure, corrupt regulators may divert a country’s investment away from the highest-value projects into potentially useless but secrecy-preserving projects. The demands of secrecy also induce government officials to maintain monopolies, to prevent entry, and to discourage innovation by outsiders in order to prevent the expansion of the ranks of the elite and preserve the secrecy of the existing corruption practices. Distinguishing these two views is best achieved by objective micro data. In macro data, corruption is measured as perception and is not directly comparable to tax rates. Yet with micro data, bribes and taxes can all be measured in monetary values and are directly comparable.

Colin Xu 321
Corruption is found to be more damaging than taxation by recent micro studies (Fisman and Svensson 2007; Cai, Fang, and Xu forthcoming). By exploiting a Ugandan firm dataset containing information on the estimated bribe payment, Fisman and Svensson (2007) study whether taxes or corruption is more damaging for firm growth. Recognizing that the bribe rate (that is, bribes over sales) is potentially endogenous, they instrument it with its industry-location average. This approach will correctly obtain the causal effects of corruption when the industry-location average of the bribe rate does not directly affect firm growth (other than just through firm bribe rates)—likely when there are no other industry-location-specific variables that directly affect firm growth. They find that both are negatively correlated with firm growth. A 1 percentage point increase in the bribe rate is associated with a reduction in firm growth of 3 percentage points, an effect about three times greater than that of taxation in their sample.

Evidence from China—a country featuring both spectacular growth and rampant corruption—also suggests that corruption tends to be quite damaging on average. Cai, Fang, and Xu (forthcoming) use a large sample of Chinese firms to examine the effects of corruption on firm performance. Since Chinese firms tend to vastly under-report bribes in surveys, the paper adopts an idea proposed by a local official acquaintance to use the entertainment and traveling costs (ETCs) of firms as a proxy of corruption. Indeed they find that ETCs are a good proxy for corruption: they are higher where you expect them to be higher (for example when government services are poor so you have to bribe for services as “grease payment”), and they are higher where tax burdens are high (so you can reduce the tax burden and enforcements through bribes). Using city-industry median ETCs as the instrument for the firm-level ETC, they find that ETCs on average have significantly negative effects on firm productivity, yet official effective tax rates (that is, total tax payments over sales) are not, on average, significantly related to productivity. The estimated effect of ETCs is causal if the city-industry median ETC does not directly affect productivity (other than through firm-level ETCs). Interestingly the negative effects of ETCs are less pronounced and can completely disappear in locations featuring particularly high tax burdens and bad government services. So there are private returns to firms in bribing governments, especially in areas with a particularly bad quality of governance. The negative effects of corruption are also less pronounced in low-income regions.16

The Role of Courts

The courts are an important institution for property rights protection. A sound legal system is a prerequisite for the Coase theorem to work its magic—for transacting parties to reach the social optimum (Coase 1960). A good court system
would sustain firms’ expectation that their contractual rights would be honored in the face of contract breaches, allowing them to commit necessary investments and to expand without worrying about contract reneging.

The legal system helps firms grow by improving the trust needed for new trans-action relationships (Johnson, McMillan, and Woodruff 2002). Using firm data from Poland, Romania, Russia, and Slovakia to examine how court quality affects business relationships, the authors find that beliefs in the effectiveness of courts are associated with a higher level of trust shown in new relationships between firms and their customers. Well-functioning courts also encourage entrepreneurs to try out new suppliers, which facilitates new entry and firm expansion.

The legal system also reduces the reluctance to expand and thus increases firm size where owners cannot effectively diversify their idiosyncratic risks. Using aggregated firm census data in 1998 and focusing on within-Mexico cross-state industry variations in firm sizes and the quality of local legal institutions, Laeven and Woodruff (2008) find that firm size increases with the quality of the local legal system. Furthermore the legal system affects firm size by reducing the idio-syncratic risk faced by firm owners—the legal system has a smaller impact on partnerships and corporations than on proprietorships, where risk is concentrated in a single owner. They deal with the potential endogeneity of the quality of the local legal system by instrumenting it with historical conditions (the share of the indigenous population in 1990 and the number of cultivated crops with large scale economies, that is, sugar, coffee, rice, and cotton in 1939). The results are also robust with respect to alternative size and institutional measures.

There is evidence that courts improve efficiency in the bankruptcy process (Gine and Love forthcoming). Bankruptcy laws are supposed to reorganize viable firms and liquidate nonviable ones to preserve efficiency. Indeed nearly 90 countries have reformed their bankruptcy systems since World War II. Yet, in developing countries, the reorganization processes are lengthy and costly. Does efficiency improve with lower reorganizing costs by reducing the statutory dead-lines for reorganization plans? Gine and Love (forthcoming) exploit the bank-ruptcy reforms in Colombia in late 1999 to examine that question. By using a unique dataset with 1,924 bankruptcy cases filed between 1996 (before the reform) and 2004 (after the reform), they obtained three findings. First, the dur-a-tion of reorganization proceedings dropped significantly. Second, under the new law, firms filing for reorganization are healthier and more viable than those filing for liquidation. Yet, under the old law, these two types of firms are similar. The new law thus better separates viable and nonviable firms and achieves the effi-ciency-enhancing goal of the bankruptcy procedure. Third, financially distressed firms recover better under the new law. One year after filing for reorganization, they are more likely to improve performance after the reform. Thus improving bankruptcy regulations allows financially distressed yet viable firms to improve.
But a good legal system is not a precondition for a country to grow (Landa 1981; McMillan and Woodruff 2002). Indeed an influential theory about the legal system, the development theory, asserts that courts work better in countries with richer and more educated populations (Demsetz 1967). A better-educated population raises both the efficiency of courts (since education is an input) and the demand for them. Courts are only worthy when transactions and contracts become more complicated and the demand for arms-length enforcement becomes sufficiently high. For poor societies, informal dispute resolutions often prove to be sufficient (Landa 1981). Consistent with this theory, McMillan and Woodruff (1999, 2002) show how entrepreneurs in transitional countries overcome imperfection in the legal systems. In Vietnam, more than 90 percent of firm managers responded that they did not rely on courts for conflict resolution. Instead they rely on ongoing relationships and the threat of losing future businesses. They often emphasize the continuation of existing relationships and punish bad behavior less severely than expected when one party breaches a contract. Without an effective court system, China substitutes family control of businesses to avoid relying on the court system (Lu and Tao 2009). All of these alternative mechanisms reduce the reliance on courts for contract enforcement.

When stakes are high, however, firms still rely more on formal contracts or advance payments, along with community sanctions (Cooter and Landa 1984). Indeed, using the World Bank Investment Climate data of China, Long (forthcoming) finds that a better local court system, as proxied by the city-industry average share of all business disputes resolved through the court, is associated with a higher investment rate, more innovation, and more complex transactions in several relatively advanced cities in the early 2000s. The results hold even with controls for city and industry fixed effects. The effects of court quality on firm outcomes are thus inferred from relating firm outcomes to the variations in court quality across cities within the same industry. Thus law seems to play a positive role even in a quintessential “poor law yet high growth” country in its more mature stage of transition for selective advanced cities (since the sample cities in Long’s study are the most advanced ones in China).

The Effects of Labor Regulations

Referring to the rules and regulations by which governments control how firms manage labor, labor regulations are tighter when firms have less discretion in freely choosing and adjusting the quantity, quality, and prices of labor. Particular types of labor regulation include employment protection through severance payments, advance notice of dismissal, administrative authorization, and prior negotiation with trade unions (Cahuc and Zylberberg 2004). The Organisation for
Economic Co-operation and Development (OECD) has an index of labor flexibility, according to which Canada, the United Kingdom, and the United States are more flexible than France, Germany, and southern European countries (Italy, Portugal, and Spain). The study of the effects of labor regulations at cross-country level is often inconclusive, partly because the data on labor regulation are available only for a limited number of industrialized countries. The estimated effects of labor protection on the level of unemployment range from positive (Lazear 1990), to negative (Nickell 1997), to insignificant (Bertola 1990). Thus the existing cross-country studies on labor regulations cannot offer much guidance for developing countries.

Recent firm-level studies of labor regulations in developing countries have substantially improved our understanding about how labor regulations work in these countries. They suggest that labor flexibility facilitates better firm performance, faster factor adjustments, and a more efficient distribution in firm sizes. China has been a good example. During the past two decades, China featured remarkable labor flexibility and economic growth (Dong and Xu 2008, 2009). Thus it is not surprising that labor flexibility is found to be efficiency-enhancing in China. Hallward-Driemeier, Wallsten, and Xu (2006) use the World Bank Investment Climate Survey in five Chinese cities to examine how the investment climate affects firm performance (that is, sales growth, investment rate, TFP, and employment growth). After controlling for firm characteristics (age, size, and ownership), city characteristics (population and income) and other investment climate indicators, including city and industry fixed effects, firm performance is still significantly better in city-industry cells that feature a higher share of nonpermanent workers. This is consistent with the notion that labor flexibility allows firms to adjust more easily to changing economic circumstances and to be more productive. A caveat is that other city-industry-specific variables may still account for the positive correlation between firm performance and our proxy of labor market flexibility. Why does labor flexibility improve firm performance in China? Facing adverse demand shocks, firms with more nonpermanent workers find it easier to adjust their labor forces and therefore to reduce costs and restore optimal factor allocations. Furthermore, firms with a flexible labor force do not have to fear labor hold-ups when considering technology and investment decisions, and the choices of technology and capital–labor ratios would thus be more efficient.

Relatedly, cumbersome labor regulations are found to be associated with smaller firm sizes and more informality in India (Amin 2009a). Using World Bank Investment Climate data on retail sectors, Amin proxies cumbersome labor regulations as the state share of firms viewing labor regulations as minor or major obstacles. He finds that this measure is robustly correlated with smaller firm sizes and informality, even after a series of sensitivity checks: controlling
for development level, store characteristics, city- and state-level variables, and a proxy of general regulation burdens. His results are also robust for small and large firms. A caveat for his study is that this regulation proxy may merely capture the effect of related state-level variables.

Facing bad labor regulations, firms do adjust on other margins. Relying on data of 2000 retail stores in India, Amin (2009b) finds that stores located in states with more cumbersome labor regulations are more likely to adopt computer technology, consistent with the notion that labor-saving technology will be adopted to cushion the blow of labor regulations. This result is robust whether they use the regulation index provided by Besley and Burgess (2004) or the share of firms in a state viewing labor regulations as “minor obstacles and above” in his data. Similarly, Adhvaryu, Chari, and Sharma (2010) use comprehensive firm-level data (aggregated to the level of district) in India to study how rainfall shocks affect labor adjustments differently in Indian states with various degree of labor regulations—again using the Besley–Burgess index of labor regulations in India. They find that, facing rain shocks, districts with more flexible labor regulations are able to adjust their labor to a greater extent. Moreover the labor adjustment effects exist only for regulated firms (that is, firms with more than 50 employees)—and do not exist for small firms that are not regulated in labor. The evidence thus points to causal effects.²¹

Recent evidence from Indonesia highlights the tradeoff between equity and efficiency associated with labor regulations. In the 1990s, Indonesia experienced two changes in labor regulations. First, the minimum wage more than doubled. Second, there was a strong antisweatshop campaign targeted at the textile, footwear, and apparel sectors, especially in those districts housing Nike, Addidas, and Reebok. As a result of the antisweatshop campaign, the targeted firms were induced to sign codes of conduct pledging to raise wages and improve working conditions. The campaign therefore amounts to informal labor regulations. Harrison (2010) use two waves of the annual manufacturing surveys of Indonesia to identify the causal effects of these two types of labor regulation through the difference-in-differences approach. By comparing the before–after difference for the treated group and the before–after difference for the comparison group (after controlling for other necessary covariates), Harrison identifies the effects of the change in minimum wage and the antisweatshop campaign on wages, employment, and other firm outcomes. She also demonstrates the robustness of key results using a variety of robustness checks such as an alternative definition of treatment and various controls of confounding factors. The results indicate that minimum wages have a significantly negative effect on employment, so there is a tradeoff between quantity and quality of jobs. In addition, the antisweatshop campaign is found to increase the wages of affected firms by 10–30 percent. While the campaign did not have additional adverse effects on
employment within the affected sectors, it led to falling profits, lower productivity growth, and plant closures for smaller exporters—so the seemingly pro-equity labor regulations reduce equity for workers of small exporters.

Stringent labor regulations have allocation consequences in high-income developing countries as well. Scoring the highest in an index of strictness of employment law (Botero and others 2004), Brazil is one of the most regulated countries regarding labor in the world. Not surprisingly, in 1999, 40 percent of the private sector was informal. Almeida (2005) investigates how regional differences in labor regulation enforcement affect informality and firms’ labor productivity. Labor regulation is measured as the number of fines related to labor issued in each region. She finds that a stricter enforcement of labor regulation leads to less informality but lower productivity and investment. The results are robust when the labor regulation proxy is instrumented by measures of access of labor inspectors to firms and measures of general law enforcement in the area where the firm is located. Using the same data, Almeida and Carneiro (2009) find that a strict enforcement of labor regulations constrains firm size and leads to higher unemployment. Recognizing that the enforcement of labor regulations may be endogenous, the authors instrument it with the distance between the city where the firm is located and the surrounding enforcement offices, while controlling for a rich set of city characteristics.

A channel through which labor regulations worsen efficiency is by increasing the discrepancy between labor costs and labor productivity and by increasing the range of productivity across firms (Petrin and Sivadasan 2006). Proxying job security regulations by the costs of dismissing employees, they investigate the effects of two changes in Chile. In 1984, the government no longer exempted firms that could show “economic cause” for dismissal from severance pay. In 1991, the government increased the ceiling of severance pay from 5 to 11 months and added a 20 percent surcharge if the employer could not prove economic causes. Petrin and Sivadasan assess the welfare effects by measuring the mean and the variance of the difference between the marginal revenue product and the marginal input prices, using the Chilean manufacturing data from 1979 to 1996. They find a substantial increase in both the mean and the variance of the within-firm gaps in response to increasing firing costs. The timings in the increasing gaps and in the regulatory changes are consistent. In contrast, the gaps do not increase for inputs that are not directly affected by firing costs.

Another channel through which stringent labor regulations hurt efficiency is by slowing down the creative destruction process, that is, the dynamic process in which inefficient firms exit accompanied by new firms entering the market and finding out about their capacity. Haltiwanger, Scarpetta, and Schweiger (2008) study how labor regulations affect job turnover by using rich new firm-level data on job flows across industries and size classes for 16 industrial and emerging
economies over the past decade. They examine whether regulations at the country level affect more regulation-vulnerable industries to a greater extent. Regulation vulnerability is proxied by the natural labor turnover rate in the corresponding industry in the United States. They find that stringent hiring and firing regulations reduce job turnover, especially in industries and size categories that inherently exhibit more job turnover.

Finally, deregulations in labor are found to facilitate factor adjustments and to enhance productivity. In the 1990s, Colombia reduced dismissal costs by 60 to 80 percent, made the social security system more portable, and minimized controls over capital market (by liberalizing foreign direct investment). Using the Colombia Annual Manufacturing Survey between 1982 and 1998, Eslava and others (2004, 2006) find that market-oriented reforms were associated with increasing employment adjustments (especially on job creations) and investments, but less with capital deployments. Furthermore market reforms are associated with rising overall productivity that is largely driven by reallocation away from low- to high-productivity businesses.

The Effects of Entry Regulations

Starting a business is costly in many countries. In Mozambique the owner must go through 19 procedures, take 149 days, and pay US$256 in fees. In Canada it takes only two procedures, two days, and US$280 (Djankov and others 2002). Combining time and out-of-pocket costs, the world average of the full cost is 66 percent of per capita GDP, varying from 1.7 percent in New Zealand to 495 percent in the Dominican Republic. Partly due to the overall change in prevailing beliefs about what is best for growth, partly due to new measurements of the costs of registering businesses across the world, entry deregulation has become a major area of reforms in the past decade.22

There is evidence that entry deregulations improve productivity and macroperformance (Loayza, Oviedo, and Serven 2005a; Crafts 2006; Barseghyan 2008). Moreover the positive effect of deregulation is found to differ by the initial level of regulation. Gorgens, Paldam, and Wurtz (2003), using the Index of Economic Freedom, find that deregulation from a high to a moderate level of regulation has a large effect on growth of about 2.5 percentage points, but further deregulation has no effects. This explains why deregulation in countries such as China and India have spectacular effects, but barely noticeable ones in OECD countries. Recent firm-level evidence of entry regulation sheds light on the specific channels through which entry regulations affect economic outcomes.

An important channel for deregulation effects is by allowing for an easier entry. Investigating the effects of entry regulations with a database of firms in
western and eastern Europe, Klapper, Laeven, and Rajan (2006) interact industry characteristics with country-level regulation indicators to examine whether regulation-vulnerable industries are more hampered by certain regulations. After controlling for country- and industry-specific factors and using the difference-in-difference approach, they find that entry regulations hamper entry, especially in industries featuring high entry (judging by what happens in the United States). Value added per worker in high-entry industries grows more slowly in countries with more onerous regulations on entry. Interestingly regulatory entry barriers do not hamper entry in corrupt countries, but do so in less corrupt ones. The results are intuitive since de facto enforcement of labor regulations in corrupt countries is less demanding. Not all regulations are bad. Regulations protecting intellectual property rights or enhancing the financial sector lead to a greater entry in industries that need more external finance or research and development. Similarly, using the United Nations Industrial Development Organization’s industry-level data in 45 countries to examine the effects of entry regulations on entry patterns across industries, Ciccone and Papaioannou (2007) find that countries featuring less registration costs see higher entry rates in industries featuring stronger global demand and faster technology changes.23

Another avenue for entry deregulation to matter is to introduce more productive (new) firms and to change sector composition. Surveying the entry effects of foreign banks, Clarke et al. (2003) show that entering foreign banks are more efficient than local banks in developing countries. Thus allowing foreign bank entry raises the overall efficiency level of the banking sector—both due to the composition effects and the rising competitive pressure for domestic firms.

Besides efficiency, entry deregulations also enhance equity by facilitating job creation. Yakovlev and Zhuravskaya (2007) examine the policy experiment in Russia between 2001 and 2004 that dramatically simplified registration and licensing procedures and reduced inspections for existing firms. The new law required a “one-stop shop” and no more than a week for registration. Each inspection agency can inspect a business no more than once in two years. Licenses are valid for no less than five years. They want to understand whether the deregulation reforms reduced regulation burdens, and whether regulation burdens affected entry, small business density, and public goods provision. They first show that the national deregulation experiments reduced regional regulation burdens using the repeated cross-sectional firm data, controlling for regional fixed effects, firm characteristics, and regional characteristics. More interestingly, they allow the deregulation effects to depend on initial regulation burdens and local institutional details (such as fiscal incentives and local government accountability). They find that the deregulation reforms significantly reduced firm regulation burdens, and the drop in regulation burdens are greater where local governments are more accountable and have stronger fiscal incentives. They then relate the measures of

Colin Xu

329
regional-level business entry, small business density, public health, and pollution to local regulation levels, controlling for regional fixed effects and time fixed effects. Since the local regulation level suffers from reverse causality—the size composition of local businesses and welfare may determine local regulation level—they instrument local regulations with the interaction of the post-deregulation dummy with the aforementioned institutional variables, holding constant regional fixed effects (since institutions as captured by the regional fixed effects directly affect the outcome variables). They find significant positive effects of deregulation for firm entry and small firm density, but not for pollution and public health.\textsuperscript{24}

The effects of entry regulations hinge critically on the implicit incentives and the specific context. For instance the effects of foreign entry depend on how close the incumbent firm is to the technological frontier, which affects the incentive to innovate (Aghion and others 2009). Aghion and others use an unbalanced panel of manufacturing firms in the United Kingdom between 1987 and 1993 to examine how foreign entry affects productivity growth and incumbent innovations. Recognizing that entry is potentially endogenous, they instrument it with policy interventions that affect the ease and cost underlying entry threat and actual entry—such as large-scale privatization, time-varying indicators of the implementation of the EU Single Market Program (SMP) in industries with medium or high entry barriers that were likely to be reduced by SMP, and competition investigations by the U.K. authority that culminated in entry-inducing remedies. They find nonuniform effects of entry on incumbent innovation—stronger for firms closer to the technological frontier. They suggest that this is due to the “escape-entry” effect: for firms sufficiently technologically advanced, incumbents work harder to innovate to win the technological race and to prevent entry or to mitigate the effects of entry. For technological laggards, the entry effect is to discourage the incumbents to innovate—they are so behind that they cannot possibly win the technological race and therefore they give up on innovations.

The importance of incentives for deregulation to work is also manifested in the telecom deregulation movement in the 1980s and 1990s, during which period national carriers were privatized, new competitors licensed, and new services allowed (Li and Xu 2002). More than 150 countries introduced new legislation or modified existing regulation. Using a comprehensive country-level panel dataset between 1990 and 2001 augmented by operator-level data on privatization and competition—and relying on the difference-in-difference approach to identify the reform effects—Li and Xu (2004) study how telecom liberalization and deregulation affect performance. They find that new entry into the sector improves both factor allocation and productivity. Most importantly, new entry and privatization are complementary in deepening network penetration and restraining the rise in service pricing.
Interestingly competition (and privatization) also reduces corruption. A study, using the World Business Environment Survey data of 21 transitional countries in East Europe and Central Asia, finds that utility employees are less likely to take bribes in countries with more competition in the utility sector and where utilities are private or privatized (Clarke and Xu 2004).

Conclusions and Policy Implications

I reach two main conclusions from the firm-level research based on the World Bank and other data related to business environments. First, some basic elements of the business environment are strongly associated with better economic performance. A basic protection of property rights from the grabbing hands of the government proves to matter a great deal for most developing countries. It has significant explanation power for firm sale growth in 54 countries. The effects of corruption are worse than those of taxes in both China and Uganda. Corruption also slows down firm entry. Thus most developing countries must contain corruption and government expropriation. Research also suggests that economists need to find out the institutional causes of corruption and to deal with it from its institutional root.

Another key ingredient of a good business environment is labor flexibility. Brazil, a middle-income country with severe labor regulations, has a larger (less productive) informal sector than its income level predicts. Lowering labor adjustment costs increased the efficiency of labor allocation in poor Colombia and rich Chile. Cross-country firm data further show that labor regulations reduce job turnover, especially in industries that are more dynamic and technologically advanced. Thus governments of developing countries, especially those with onerous labor regulations, should examine how their labor regulations compare with other countries and whether they can be relaxed to facilitate growth.

For most industries and countries, entry deregulation appears to be a good idea. In Russia and Mexico, entry deregulations create jobs and reduce prices and incumbent profits by stronger competition. Telecom deregulations around the world improve both factor allocation and productivity. Foreign entry in the United Kingdom facilitated technological innovation for firms close to international production frontiers. Competition reduces corruption. The policy implication is that without special concerns, the entry to an industry should not be heavily regulated. For industries with strong entry regulation, one should examine whether such regulations are hindering the growth and innovation of their firms. Foreign entry into sectors that are close to the international production frontier are especially encouraged.

There is some evidence that efficient legal systems for facilitating exits have high payoffs for developing countries. Reducing reorganization processes by
reducing statutory deadlines appears to have large payoffs—reducing time costs, attracting more viable firms to go through the process, and recovering better. Since legal procedures for bankruptcies are very costly in developing countries, especially for those featuring strong legal formalism, such countries should aim to reduce legal formalism for bankruptcies. For instance they could adopt foreclosures with no or limited court oversight and floating charges (that is, transfer control of the firm to the secured creditor).

Second, the effects of many elements of the business environment depend on industry, complementary institutions, and the initial business environment. For instance the effects of contracting institutions (such as access to finance and courts) appear to become stronger as an economy becomes more mature. In the early stage of development and transition, substitution institutions, such as clustering, reputation mechanisms, relationship contracts, and informal trade credit, could be sufficient to induce economic growth. This is found in China, Vietnam, and the early transitional eastern European countries. However, as transition and development move along, the extent of the market becomes larger, transactions become more complicated, and contracting relationships based on personal ties become insufficient. Now formal market-supporting institutions are needed to encourage arm’s-length contracting. While property rights institutions are found to be much more important than access to finance for European and Central Asian countries in early periods of transition, access to finance becomes visibly more important for China at a later stage of development, although property rights institutions remain important. In addition there is also evidence that courts have become important for eastern European transitional countries and even in advanced parts of China. Furthermore courts are found to be especially important when human capital looms large, likely in industrialized countries such as in western Europe. The policy implication is that, as they develop further, governments should build and improve contracting institutions, such as formal finance and courts, as they can better support the transactional needs of larger and more sophisticated firms. Governments that lag behind in facilitating such institutional changes would slow down development.

Infrastructure appears to be particularly important for poor developing countries. In a sample of poor countries (Bangladesh, China, Ethiopia, and Pakistan), infrastructure has positive effects for productivity, factor returns, and international integration. Yet there are no significant effects across regions within China around 2002. Thus infrastructure investment seems to have higher returns in countries with poorer infrastructure.

The effects of entry deregulation differ by industries. Entry deregulations have particularly pronounced effects in industries with natural high-entry rates. For industries heavily dependent on fresh ideas, such as IT and R&D intensive industries, entry deregulations are therefore particularly important. Even though
allowing for easier entry appears to be a good idea for most industries, some regulations remain useful. Regulations that protect intellectual property rights or enhance the financial sector facilitate more entry into R&D intensive and finance-dependent industries, which tend to be high value-added industries enabling sustainable growth for middle- and high-income countries.

Entry effects also differ by initial conditions. The benefits of opening up entry for foreign firms seem stronger when domestic firms are competitive and face strong incentives. The effects of foreign entry are stronger for firms closer to technological frontiers due to the escape-entry effects—competitive domestic firms can potentially win technology races and therefore have stronger incentives to compete. For domestic firms that are completely behind in the technology race, the innovation effects are low since they give up without hope of ever winning. Furthermore there is also evidence that entry deregulations are more effective when coupled with privatization if the incumbents are state owned. The policy implications are that countries can target entry such that frontier industries reduce entry barriers for foreign firms. They should adopt complementary policies to facilitate resources shifting to more advanced industries where incumbents react more positively to entry threat and should discourage regulations hindering such reallocations (such as through subsidies for nonviable industries). In countries with state-owned firms, privatizing them before opening up to foreign competition may be useful.

The conclusions reached in this survey are tentative. While there are many studies based on the difference-in-difference and the instrumental-variable approach that are more plausible in establishing causality, the rest of the studies are based on cross-sectional correlations, making causality inference very difficult. There are also clearly alternative explanations for this body of new findings. Since it is possible that some omitted variables account for the correlations of our proxies of business environment with economic outcomes, the conclusions I present here should be viewed with caution.

Notes

Lixin Colin Xu is a lead economist at the World Bank; Mail Stop, MC 3-307, 1818 H Street, N.W., Washington, DC 20433; email address: lxu1@worldbank.org. I am grateful to Manny Jimenez, Mohammad Amin, Robert Cull, Ann Harrison, Guofang Huang, Justin Lin, Cheryl Long, Rita Ramalho, Yan Wang, and especially three anonymous referees for helpful discussions, comments, and suggestions. The opinions expressed here are the author’s alone and do not implicate the World Bank, its executive directors, or the countries that it represents.

1. The terms ‘investigation climate’ and ‘business environment’ are used interchangeably in this paper.

2. Other features, such as geography and weather, also belong here, but since little can be done to alter them their effects are not discussed in this paper.
3. The touting of micro data here should not be interpreted as discriminating against studies based on sector-or country-level data, which are complementary to those on firm-level data. Some key explanatory variables differ only at the macrolevel, in which case it is only natural to rely on such data for identification. Often studies based on macrodata are useful first steps in our quest for understanding a specific topic. A good example is that Knack and Keefer (1995, 1997) rely on macro indicators to show the potential importance of institutions and social capital, which then usher in numerous microstudies to examine the issues more closely.

4. Note that while the de jure institutions are held constant within a country, the de facto enforcements of institutions across regions are not (Hallward-Dreimer, Khun-Jush, and Pritchett 2010). Only to the extent that de jure institutions capture a significant source of variations for de facto institutions, relying on within-country variations mitigate the omitted variable bias due to the lack of control for institutions.

5. The World Bank Enterprise Surveys, including questionnaires, are available at http://www.enterprisesurveys.org/.


7. For the use of the difference-in-difference approach in the business environment literature, see Klapper, Lueven, and Rajan (2006), Ciccone and Papaioannou (2007), and Aghion and others (2008).

8. Shleifer also considers state ownership, which is beyond the scope of this paper and is therefore ignored.

9. Facing specific bottlenecks, countries have found unique ways to make things work. China, for instance, was jump-started by leasing land to farmers and by letting them decide what to do with their lands (Lin 1992; Lin, Cai, and Li 2003), by encouraging entry through the emergence of township-and-village enterprises (for example semigovernment, semiprivate firms with hard budget constraints) and by giving state-owned enterprises autonomy and incentives to their operations (Li 1997; Xu 2000; Lin, Cai, and Li 2003). In Vietnam, court enforcements for new entrants were facilitated by reputation-based informal contracts without the standard court system (McMillan and Woodruff 2002).

10. The positive relationship between investment and business environment in general and infrastructure in particular does not have to hold everywhere—there could be a negative relationship in some small neighborhood (or on a more grand scale in some former socialist countries such as the Soviet Bloc). An example is that infrastructure is much worse in India than in China. To adapt to the bad electricity system, firms tend to purchase their own power generators, which increases investment. So this example may run counter to the positive relationship between investment and infrastructure. Even though Indian firms may have more power generators, they still have lower capital–labor ratios relative to China, which possesses a better infrastructure (based on the author’s check on investment climate data for both countries).

11. In a similar vein, Dollar and Kraay (2003) find that once one allows the same set of factors determining both trade and institutions and properly deals with the endogeneity of both factors, institutions no longer play a major role in explaining growth, though trade does.

12. Similarly, using the World Bank investment climate survey, Lu, Png, and Tao (2008) find that better institutions (that is, those with property rights protection and contract enforcement) facilitate firm growth in China. This is true even after they instrument institutions with city population density in 1918–19 and whether the city was administered by the British in the late Qing Dynasty. While this study strengthens the finding about the importance of institutions for firm growth, its instruments are not fully convincing—one can imagine that the historical variables might affect firm growth via routes other than institutions, for instance, through better infrastructure.

13. A caveat to this study is that other country-level variables related to property rights can account for the correlation between the measures of property rights and the access to external finance.
14. The property rights measures are broad cross-country indicators (such as those from ICRG, the Heritage Foundation, the U.S. Trade Representative, and so on).

15. See the earlier discussions on the details of measurements.

16. Relatedly, firm performances in China are significantly lower in locations and industries with a higher share of informal payment to sales for firms (Hallward-Driemeier, Wallsten, and Xu 2006). The effects of corruption in China are likely underestimated due to the attenuation bias of measurement errors—the incidence of corruption in the data is far too low to be plausible.

17. Labor regulations have other important objectives, such as health, job safety, job security, and reducing income inequality. Since the new empirical literature on labor regulations is mainly concerned with the income and growth effects, I do not have much to say about these additional objectives.

18. Indeed around the turn of the century, China underwent the most dramatic labor-restructuring program in the world (Dong and Xu 2008, 2009). Between 1995 and 2001, more than 35 million state workers had been laid off. In examining the enterprise restructuring in the early 2000s, the authors found that firms were more likely to undergo downsizing when they were state-owned enterprises (SOEs); when these SOEs were older, larger, and had more excess capacity; and when product prices dropped, indicating that the labor restructuring tends to be efficiency enhancing. Moreover the patterns of labor adjustments for private enterprises and SOEs were similar (such as reducing labor demand when output prices drop and when wages increase), indicating an increasingly integrated Chinese labor market.

19. Fernandes and Pakes (2008) also find evidence of labor misallocation in India by using the World Bank Investment Climate Survey data. They find that the growth rate of the Indian manufacturing sector is significantly lower than its growth rate in the services sector or in China. They compute the ratio of labor under-utilization and find that this is widespread in Indian states, especially in those states with lower income levels and for productive firms.

20. Policies that increase the informal sector likely reduce the average productivity and income of the economy. Recent evidence suggests that the informal sector tends to attract low-productivity workers who have a low likelihood of ever being promoted to the formal sector (Maloney 2004; de Mel, McKenzie, and Woodruff 2008; La Porta and Shleifer 2008; Bruhn forthcoming). The most important determinants of the size of the informal sector across countries include the stringency of labor regulations and their enforcement (Johnson, Kaufman, and Zoido-Lobaton 1998; Loayza, Oviedo, and Serven 2005b).

21. Surprisingly they find that firms in districts with more cumbersome labor regulations are able to adjust their other production factors (materials, fuel, and capital) to a greater extent—such that, when examining the effects of labor regulations on total value added and profits, there are no differences between states with heavy and those with light labor regulations.

22. Djankov (2009) surveys how entry deregulation has affected economic performance over the past decade.

23. The entry effects of deregulation can sometimes be achieved without explicit entry deregulation. Long and Zhang (2009) find that clustering (that is allowing firms producing different parts of the same product to be located together) reduces financing needs, increases entry and competition, and therefore improves export and productivity.

24. Similar benefits are found in the deregulation experiment in Mexico by Bruhn (forthcoming), which uses cross-municipality and time variations in implementation to study the effects of reducing registering costs on the number of firms, employment, prices, and income. Mexican registration reforms reduced registration procedures from eight to three on average. Relying primarily on the difference-in-difference approach to infer the reform effects, the author finds that the reform increases the number of registered businesses by 5 percent in eligible industries. It creates more jobs: employment in eligible industries increase by 2.8 percent after the reform, benefiting primarily those previously unemployed and out of the labor force. It also benefits consumers at the expenses of the incumbent business owners: prices drop by 0.6 percent after the reform, while the increased competition associated with more entry reduces the average income of incumbent owners.
by 3.2 percent. The same policy experiment is also studied by Kaplan, Piedra, and Seira (2007), who find that new start-ups increased by around 4 percent for affected industries.

References

The word processed describes informally reproduced works that may not be commonly available through libraries.


336


