Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?

JAMES R. TYBOUT

1. Overview

The manufacturing sector is often the darling of policy makers in less developed countries (LDCs). It is viewed as the leading edge of modernization and skilled job creation, as well as a fundamental source of various positive spillovers. Accordingly, although many LDCs have scaled back trade barriers over the past twenty years, the industrial sector remains relatively protected in the typical country (Maurice Schiff and Alberto Valdez 1992, ch. 2; Refik Erzan et al. 1989; Francis Ng 1996). Governments also promote manufacturing with special tax concessions and relatively low tariff rates for importers of manufacturing machinery and equipment.

At the same time, many observers believe that the maze of business regula-

1 Pennsylvania State University. This paper was partially funded by the World Bank’s Economic Development Institute. I have benefited from discussions with Bee-Yan Aw, Mark Dutz, Mark Gersovitz, Bernard Hoekman, Carl Liedholm, William Maloney, Desmond McCarthy, Mark Roberts, Kamal Saggi, Maurice Schiff, and William Steel. I am especially indebted to John Pencavel, John McMillan, Howard Pack, and several anonymous reviewers, who have improved this manuscript with many helpful suggestions.

2 The need for revenue is a second motivation for relatively high tariffs in developing countries, although nontariff barriers seldom serve this function.

3 A well-known example of the problem was generated by the Institute for Liberty and Democracy in Peru, which attempted to register a fictitious clothing factory in the mid-1980s. “To register the imaginary factory took 289 days and required the full-time labor of the group assigned to the task, as well as . . . the equivalent of 23 minimum monthly wages” (Hernando de Soto 1989, p. xiv).

4 For purposes of this paper I will ignore the fact that Turkey, Mexico, and Korea are members of the OECD.

5 There were no major differences between the LDCs and the OECD in terms of the regulations
Moreover, within the manufacturing sector, it is also often argued that policies favor large firms while inhibiting growth among small firms (Ian Little 1987). In some cases, investment incentives are available only to projects above a minimum scale, and large-scale producers are singled out for special subsidies. Anti-trust enforcement is typically weak, and special tax breaks are sometimes meted out to large, influential corporations (Bernard Gauthier and Mark Gersovitz 1997).

Even when policies do not explicitly favor large firms, these firms may enjoy de facto advantages. Banks view them as relatively low risk and cheap to service (per unit of funds lent), so they have preferential access to credit. This phenomenon is relatively marked in developing countries because private sector credit is relatively scarce there, information networks are poorly developed, and binding interest controls are relatively common (Ross Levine 1997; Little 1987; Tybout 1984).

Protectionist trade regimes are also more likely to favor large firms, both because these firms’ products compete more directly with imports, and because sectors with large, capital-intensive firms lobby the government more effectively. Further, while many of the costs of dealing with dense regulatory regimes are fixed, the payoffs from doing so probably increase with the scale of operations. For example, access to the legal system, access to the formal banking sector, and publicly administered employee benefits are relatively valuable to large producers (Alec Levenson and William Maloney 1997). Hence, moderate-sized firms, which are large enough to show up on regulators’ radar screens but too small to gain much from compliance, may be punished the most.

These basic tendencies of LDCs—toward industrial sector promotion, dense, unpredictable regulatory regimes, and credit markets or commercial policies that favor large firms—raise a number of fundamental empirical issues. First, do the regulatory regimes and the bias against small producers prevent small firms from growing, and thereby create losses due to unexploited scale economies? Second, if these regimes prevent small firms from threatening the larger incumbents, do LDC industrial sectors lack dynamism and competition? That is, have entrenched oligopolies emerged that are neither innovative, technically efficient, nor likely to price competitively? Finally, has trade protection compounded the technical inefficiencies and monopoly power that arise from regulatory regimes? In this paper I selectively take stock of what we have learned about these issues from firm- and plant-level data sets over the past twenty years.

I shall begin by briefly reviewing some of the distinctive features of the environment in which LDC manufacturers operate. This will serve as background for the discussion that follows, and help to distinguish those features of

---

6 See Howard Pack and Larry Westphal (1986) on Korea; Marluz Cortes, Albert Berry, and Ashfaq Ishaq (1987) on Colombia; Robert Wade (1990) and Bruch and Hiemenz (1984) on E. Asia. India is an exception—see, for example, Little, Mazumdar, and Page (1987).

7 Levine (1997) provides references.
LDC manufacturers that trace to structural characteristics of their economies from those induced by industrial, trade, or labor policies. Next, drawing on the available evidence, I will take up the issue of whether small firms have been somehow suppressed, and more generally, whether the LDC business environment has bred noncompetitive pricing behavior and low productivity. Finally, I will address the question of how trade protection has conditioned pricing, efficiency, and productivity growth.

2. The Business Environment: What’s Different in LDCs?

In terms of income levels and business environment, the countries typically labeled “developing” are a very heterogeneous group. By the World Bank’s reckoning they currently span the per capita income range from $US 80 (Mozambique) to $US 8,380 (Argentina). Nonetheless, looking across countries, some distinctive features of the business environment become increasingly evident as one moves down the per capita income scale. At the risk of over-simplifying, I will begin by mentioning the most striking and widely acknowledged among them.

Market size. Although some developing economies are quite large, most are not. Hence, excepting countries like

8 This income range describes all countries that are not classified by the World Bank as high income or transition. Figures are taken from the World Bank’s World Development Indicators 1998.
Brazil, China, India, and Indonesia, the size of the domestic market for manufactured products is relatively limited (Figure 1). Further, among the least developed countries, Engel effects favor basic subsistence needs over all but the most basic manufactured products (Figure 2). So when transport costs are significant and the OECD countries are distant, demand for the more sophisticated manufactured goods is small.

**Access to manufactured inputs.** The menu of domestically produced intermediate inputs and capital equipment is also often limited in developing countries. Thus producers who might easily have acquired specialized inputs if they were operating in an OECD country must either make do with imperfect substitutes or import the needed inputs at extra expense. This latter option is the dominant choice among smaller countries.\(^9\)

**Human capital.** Low rates of secondary education...

\(^9\)As a crude exercise, one can designate “non-electrical equipment,” “electrical machinery” and “transport equipment” as the machinery and equipment industries. Then using domestic production data from the World Bank and trade data from the COMTRADE system, the ratio of net imports to domestic consumption for this group of products can be constructed product by product. For the 53 countries with complete data, 75 percent of the cross-country variation in this measure is explained by the logarithm of GDP and a dummy for transition economies. The coefficient on the logarithm of GDP is –0.14 and it carries a t-ratio of 9.69. (Surprisingly, the log of GDP per capita adds no additional explanatory power to the regression.) The predicted share of imports in domestic consumption of machinery and equipment is greater than 0.6 for economies with GDPs less than 12.5 billion 1987 US dollars.
education and a scarcity of technicians and scientists also affect the mix of goods manufactured and the factor proportions used to produce them.\textsuperscript{10} Similarly, many have argued that flexibility in production processes and the ability to absorb new technologies are directly related to the stock of human capital (e.g., Richard Nelson and Edmund Phelps 1966; Robert Evenson and Westphal 1995; Wolfgang Keller 1996).

Infrastructure. Roads, ports, airports, communication facilities, power, and safe water access tend to be relatively limited in LDCs (World Bank 1994, p. 3, Figure 1; p. 13, Table 1.1, and p. 14, Figure 1.1). Production techniques are directly affected, as are the costs of servicing distant markets. Poor transportation networks are particularly limiting in the least developed, more agrarian economies, where consumers are spread throughout the countryside. In instances where infrastructure services are missing or unreliable, some firms must produce their own power, transport, and/or communication services.

Volatility. Macroeconomic and relative price volatility is typically more extreme in developing countries. Historically, Latin America and sub-Saharan Africa have stood out as the most volatile, but all developing regions have done worse than the industrialized countries (World Bank 1993; Ricardo Hausmann and Micheal Gavin 1996).

Governance. Finally, legal systems and crime prevention are also relatively poor in developing countries, and corruption is often a serious problem (World Bank 1997; Brunetti, Kisunko, and Weder 1997). Hence the protection of property rights and contract enforcement can be problematic.

3. Plant Sizes and Scale Efficiency

Combined with industrial sector policies, the above circumstances (and others I have neglected to mention) lead to several distinctive features of LDC manufacturing sectors. Perhaps the most striking of these is their dualism. In many industries, large numbers of microenterprises and a handful of modern, large-scale factories produce similar products side by side.\textsuperscript{11} The small producers frequently operate partly or wholly outside the realm of government regulation, and rely heavily on informal credit markets and internal funds for finance. They are relatively labor intensive, so they account for a larger share of employment than of output.

3.1 The Size Distribution

The contrast between the size distribution of plants in developing countries and that found in the OECD is dramatic. Table 1 provides some crude comparisons. Note that there is a large spike in the size distribution for the size class 1–4 workers, and it drops off quickly in the 10–49 category among the poorest countries.\textsuperscript{12} This is not true

\textsuperscript{10} The wages of scientists and engineers in manufacturing firms constitute 0.2 percent of GDP in the most technologically primitive of the developing countries, while they account for 1.0 percent of GDP in the OECD (Evenson and Westphal 1995, Table 37.1). A logarithmic regression of the secondary school enrollment rate on GDP per capita yields an elasticity of 0.62 and an R-squared of 0.65. (Calculations are based on the Barro-Lee data base, available at www.worldbank.org/growth/ddbarle2.htm.)

\textsuperscript{11} In Colombia, for example, more than half of the 5-digit ISIC industrial sectors that contain plants with less than ten workers also contain plants with more than two hundred workers. I base this statement on Colombia’s annual manufacturing survey, which neglects most plants with less than five workers, so this statistic substantially understates the prevalence of microenterprises.

\textsuperscript{12} Many microenterprises are invisible to official census takers because they do not have postal boxes, are impermanent, and/or are part of farm compounds. Thus Table 1 substantially understates their prevalence. "[C]omparison of village by village enterprise censuses conducted by [Michigan State University] and local scholars with
in the United States or other industrialized countries. The emphasis on small-scale production not only correlates negatively with per capita income levels across countries (Carl Liedholm and Donald Mead 1987, p. 16; Ranadev Banerji 1978), but also within countries through time (Ian Little, Dipak Mazumdar, and John M. Page, Jr. 1987; William Steel 1993).

What accounts for this phenomenon? For some LDCs, observers have pointed to regulations and taxes that are enforced only among the large, formal sector firms. James Rauch (1991) formalizes this explanation by extending Robert Lucas’s (1978) model of the firm size distribution with heterogeneous worker/entrepreneurs. He shows that when larger firms face higher unit input costs, the most talented entrepreneurs operate big firms to exploit their productivity advantage, and the extra profits they earn from being big more than cover the higher input costs they must pay. Less talented entrepreneurs stay small and informal. The size distribution exhibits a “missing middle” because it

<table>
<thead>
<tr>
<th></th>
<th>Numbers of Workers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1–4</td>
</tr>
<tr>
<td>United States, 1992(^a)</td>
<td>1.3</td>
</tr>
<tr>
<td>Mexico, 1993(^b)</td>
<td>13.8</td>
</tr>
<tr>
<td>Indonesia, 1986(^c)</td>
<td>44.2</td>
</tr>
<tr>
<td>S. Korea, 1973(^d)</td>
<td>7.9</td>
</tr>
<tr>
<td>S. Korea, 1988(^e)</td>
<td></td>
</tr>
<tr>
<td>Taiwan, 1971(^f)</td>
<td></td>
</tr>
<tr>
<td>Taiwan, 1986(^g)</td>
<td>20</td>
</tr>
<tr>
<td>India, 1971(^h)</td>
<td>42</td>
</tr>
<tr>
<td>Tanzania, 1967(^i)</td>
<td>56</td>
</tr>
<tr>
<td>Ghana, 1970(^j)</td>
<td>84</td>
</tr>
<tr>
<td>Kenya, 1969(^k)</td>
<td>49</td>
</tr>
<tr>
<td>Sierra Leone, 1974(^l)</td>
<td>90</td>
</tr>
<tr>
<td>Indonesia, 1977(^m)</td>
<td>77</td>
</tr>
<tr>
<td>Zambia, 1985(^n)</td>
<td>83</td>
</tr>
<tr>
<td>Honduras, 1979(^o)</td>
<td>68</td>
</tr>
<tr>
<td>Thailand, 1978(^p)</td>
<td>58</td>
</tr>
<tr>
<td>Philippines, 1974(^q)</td>
<td>66</td>
</tr>
<tr>
<td>Nigeria, 1972(^r)</td>
<td>59</td>
</tr>
<tr>
<td>Jamaica, 1978(^s)</td>
<td>35</td>
</tr>
<tr>
<td>Colombia, 1973(^t)</td>
<td>52</td>
</tr>
<tr>
<td>Korea, 1975(^u)</td>
<td>40</td>
</tr>
</tbody>
</table>

**Sources:**

\(^a\) 1992 U.S. Census of Manufacturing, unpublished Census Bureau calculations.

\(^b\) INEGI (1995).

\(^c\) Steel (1993).

\(^d\) Little, Mazumdar, and Page (1987, Table 6.5).

\(^e\) 1988 Census of Manufacturing, Republic of Korea, calculations of Bee-Yan Aw.

\(^f\) Chen (1997, Table 2.2).

\(^g\) Liedholm and Mead (1987).

\(^h\) Official censuses shows that the latter not infrequently underestimated the number of enterprises by a factor of two or more (Carl Liedholm and Donald Mead 1987, p. 20).

\(^i\) Small-scale production not only correlates negatively with per capita income levels across countries (Carl Liedholm and Donald Mead 1987, p. 16; Ranadev Banerji 1978), but also within countries through time (Ian Little, Dipak Mazumdar, and John M. Page, Jr. 1987; William Steel 1993).
never pays to be just large enough to attract enforcement. Related results can be obtained using dynamic industrial evolution models with entry costs for formal sector participation, as I will argue in Section 4 below.

Descriptive country studies support and elaborate upon this basic story. Writing on Peru in the 1980s, de Soto (1989) argues that many entrepreneurs remained small to avoid excessive regulation, and thus failed to challenge entrenched large firms, who moved relatively easily in the regulatory maze. On Cameroon in the early 1990s, Gauthier and Gersovitz (1997) show that small firms remained informal and avoided taxes, while large firms were influential enough to obtain special treatment. Mid-sized firms bore the highest tax burden. On India, which is unusual in the favoritism it has shown to small firms, Little, Mazumdar, and Page (1987, p. 32) write: “Not only would small firms [that graduate] have to cope with a much more difficult licensing policy, but they would also have to contend with higher labor costs (including wages and fringe benefits as laid down by labor laws) and substantially higher excise duties.”

These arguments help to explain the size distribution in those developing countries with heavy regulation. But the pervasiveness of small firms in LDCs suggests that other more universal forces are at work. For example, as noted in Section 2, the poorest countries tend to be the least urbanized, and their transportation networks tend to be underdeveloped. So small, diffuse pockets of demand lead to small-scale, localized production.  

13 Liedholm and Mead (1987) report that “in the four survey countries where relevant data were collected, direct sales to final consumers dominated [sales to businesses, government sales and exports], and, in fact, exceeded 80 percent in three of the countries.” (pp. 46–47)

Underdevelopment also spawns small firms because Engel effects skew demand for manufactured products toward simple items like baked goods, apparel, footwear, metal products, and furniture. All of these products can be efficiently produced using cottage technologies, so there is little incentive to consolidate production in several large plants and incur the extra distribution costs.

Further, plentiful unskilled labor and the lack of long-term finance create incentives to economize on fixed capital. Since most machinery and equipment must be imported, the trade regime and the lack of local technical support may further militate against factory production in small markets. 14 In the presence of wage rigidities, abundant labor and scarce capital can also mean that formal sector jobs are rationed, hence workers unable to find employment in the formal sector may create their own microenterprises to survive. 15

Finally, volatility in the business environment—both regulatory and macroeconomic—can discourage mass production techniques. Investments in fixed capital involve long-term commitments to particular products and production

14 Cortes, Berry, and Ishaq (1987, pp. 153–54) note that “the increasing availability of skill machine operators [in Colombia] has also contributed to the establishment of local importers and recon- structors of used equipment . . . ”

15 The evidence on this effect is mixed. Liedholm and Mead (1995) argue that it is important in sub-Saharan Africa, while Levenson and Maloney (1997) and Wendy Cunningham and Maloney (1998) downplay its significance in Mexico.
volumes. If there is substantial uncertainty about future demand conditions for these products, it often makes sense to choose production techniques that do not lock one in; that is, to rely more heavily on labor (Val Lambson 1991; Brunetti, Kisunko, and Weder 1997).

3.2 Are Small Firms Scale Efficient?

Does the preponderance of small firms imply that scale inefficiency is a serious problem in developing countries? Many have argued that it does, particularly in the simulation literature, where analysts often assume that the ratio of average to marginal cost is above 1.10 for the typical plant.\(^\text{16}\) However, survey-based evidence suggests that the potential efficiency gains from increases in plant size—induced, for example, by trade liberalization—are probably much smaller than these studies suggest.

The simplest studies of microenterprises relate output per worker and output per unit capital to plant size. In addition to the shortcomings of partial productivity measures, this literature suffers from several data problems. One is that the boundaries of very small firms are often ill-defined because they are part of a household or a farm, or because they are vertically integrated with nonmanufacturing activities. Another problem, stressed by Little (1987), is that many of these studies pool data on plants producing a diverse range of products. It is worth noting, however, that when Little, Mazumdar, and Page (1987) focus on four narrowly-defined Indian industries, they find “it is difficult to detect any systematic variation in labor or capital productivity with firm size.” (p. 186)

Conceptually, studies of microenterprises that attempt multi-factor productivity measures are more appealing. One strand of this literature is based on social cost–benefit ratios, constructed as the cost of labor and capital at shadow prices, relative to value-added in world prices. As discussed by Leidholm and Mead (1987), these studies have differed in their conclusions, with some finding that small enterprises are at least as efficient as others, and others finding their efficiency relatively low.\(^\text{17}\) As for the very small, Liedholm and Mead (1987) do find that one-person establishments are systematically less efficient than others, perhaps because many are created as occupations of last resort for those who cannot find work in the job market.

Scale economies are more consistently missing in studies of microenterprises based on estimated production functions. Little, Mazumdar, and Page (1987) and K. V. Ramaswamy (1994) fit simple functions to cross-sectional data on small-scale Indian producers, and report returns to scale very close to unity in all of the industries they treat. Hal Hill and K. P. Kalijaran (1993) obtain analogous results among small-scale Indonesian garment producers. Similarly, using firm-level African data collected by the Regional Program on Enterprise Development (RPED), Tyler


\(^{17}\) Leidholm and Mead (1987) find that small enterprises in Sierra Leone, Honduras, and Jamaica are at least as efficient as others. On the other hand, Sam P. S. Ho (1980) and Cortes, Berry, and Ishaq (1987) find some evidence of scale economies in Korea and Colombia, respectively. Small enterprises are typically less capital intensive, so one reason for the discrepancy may be that Leidholm and Mead use a rather high shadow price of 20 percent per annum for capital services.
Biggs, Manju Shah, and Pradeep Srivasava (1995) fit the same estimator to four manufacturing sectors in Ghana, Kenya, and Zimbabwe. Interestingly, even when the sample is limited to firms with three to twenty workers, they estimate returns to scale very close to unity. And when the entire stratified sample is used for each industry (covering the entire size spectrum), returns to scale are still close to unity in food and textiles/garments, while mild increasing returns are found in wood products and metal products.\(^{18}\)

Finally, because data sets often do not cover the smallest plants, many analysts have econometrically estimated production functions using data on plants with at least ten workers. Their dominant finding is constant or mildly increasing returns (between 1.05 and 1.10) in the various manufacturing sectors of Latin American, Asian, and North African countries.\(^ {19}\)

All of the studies mentioned in this section are plagued by measurement error problems, omitted variables, aggregation bias, and simultaneity bias (Tybout and Westbrook 1996; James Levinsohn and Amil Petrin 1997; Tybout 1992a). Nonetheless, their basic message seems consistent with engineering studies: the efficiency costs of being small are not crippling—if present at all—once the one-worker threshold has been traversed. Put differently, small firms in developing countries tend not to locate in those industries where they would be at a substantial cost disadvantage relative to larger incumbents.

### 4. Turnover, Market Share Reallocations, and Efficiency

Even if the potential gains from scale economy exploitation are small, one might argue that the prominence of small-scale producers in LDCs is symptomatic of other problems. For example, if excessive taxation and regulation keep many firms small and informal, these policies may be stanching the selection process through which better managers and/or technologies gain market share.\(^ {20}\) Severance laws and restrictions on the use of temporary workers may also inhibit the expansion and contraction of formal sector plants, limiting competitive pressures. Similarly, producer turnover may be dampened by policies that prop up "sick" firms, thereby saturating the market with inefficient producers, and discouraging better firms from entering.\(^ {21}\) Poorly functioning credit markets may further constrain entry and expansion.

#### 4.1 Analytical Models of Industrial Evolution

What might constitute evidence on these relatively subtle effects? Dynamic models of industrial evolution provide

---

\(^{18}\) This is all the more remarkable when one considers that inherently inefficient firms tend to stay small, so even in the absence of scale economies the data should exhibit some correlation between size and productivity due to selection effects (e.g., G. Steven Olley and Ariel Pakes 1996).


\(^{20}\) Of course, taxation and regulation are not inefficient per se. As Levenson and Maloney (1997) note, firms that register with tax authorities and regulators also enjoy the benefits of enforceable contracts, better access to credit, and—in the form of publicly administered fringe benefits for workers—access to risk-pooling mechanisms.

\(^{21}\) Pursell (1990) notes that "sick" enterprises propped up by the Indian government tied up roughly 14 percent of total bank credit to industry in 1986. Fikkert and Hassan (1996) review the various licensing requirements and approval procedures for capacity expansion that have prevailed in India, and provide further references.
some guidance. These models generally include representations of the processes that generate each firm’s entry, exit, productivity growth, and market share or factor use. In most modern treatments, each dimension of performance is depicted as the optimal behavior of forward-looking entrepreneurs with rational expectations but limited information.22

The literature is complex, but Hugo Hopenhayn (1992) provides a relatively tractable formulation. In his model, firms differ only in terms of their productivity levels, each of which evolves according to an exogenous Markov process. New firms enter when the distribution from which they draw their initial productivity level is sufficiently favorable that their expected future profit stream, net of fixed costs, will cover the sunk costs of entry. Firms exit when they experience a series of adverse productivity shocks, driving their expected future operating profits sufficiently low that exit is their least costly option. All firms are price takers, but the prices of their inputs and outputs depend upon the number of active firms and their productivity levels.

This model shares a number of implications with other representations of industrial evolution developed by Boyan Jovanovic (1982) and Richard Ericson and Ariel Pakes (1995). At any point in time, an entire distribution of firms with different sizes, ages, and productivity levels coexists, and simultaneous entry and exit is the norm. Young firms have not yet survived a shakedown process, so they tend to be smaller and to exit more frequently. Large firms are the most efficient, on average, so their mark-ups are the largest. Nonetheless, despite all the heterogeneity, equilibria in both Jovanovic’s and Hopenhayn’s models maximize the net discounted value of social surplus. Thus market interventions—like artificial entry barriers, severance laws, or policies that prop up dying firms—generally make matters worse.23

Under certain regularity conditions, as Hopenhayn shows, an increase in the sunk costs of entry protects incumbent firms from the upward pressure on input prices and the downward pressure on output prices that new entrants create. Thus high entry costs not only reduce the amount of entry, they encourage incumbents with relatively low productivity to stick around, and thereby increase the amount of productivity dispersion among active firms.24 In addition, the market shares of the largest, most efficient firms rise with entry costs, skewing the size distribution (Hopenhayn 1992, p. 1142). The shares of the largest firms also respond negatively to market size, since an outward shift in demand leaves the plant size density function and the underlying entry/exit processes unchanged.25

Policies that inhibit expansion or contraction have similar consequences. Using a variant on the model described above, Hopenhayn and Richard Rogerson (1993) simulate the effects of severance

22 Nelson and Sidney Winters (1982) argue that managers do not have the knowledge or the time to solve stochastic dynamic optimization problems, so these authors model entry, growth, and exit as deriving from rules of thumb that managers follow. The assumption of relentlessly optimal behavior is doubtless a caricature of the real world, but it is not obvious that alternative representations of behavior are more defensible. I will thus couch my discussion in terms of the optimizing literature.

23 Product markets are not perfectly competitive in Ericson and Pakes’s (1995) formulation so this statement does not hold for their model.

24 Exit costs have qualitatively similar effects to those of sunk entry costs because they reduce the amount of one’s initial investment that can be recovered by quitting the industry.

25 For example, if one doubles demand, concentration ratios drop by a factor of 2 and Herfindahl indices drop by a factor of 4, but turnover rates and the market shares of each quantile in the size distribution remain unaffected.
laws. They find that increases in the rate at which laid-off workers must be compensated raise the degree of persistence in firms’ market shares, increase average firm size, increase price-cost mark-ups, reduce average productivity, and reduce the job turnover rate.

If we think of sunk costs as deriving from formal sector entry rather than the creation of a microenterprise, the results I have reviewed above provide a crude basis for inference. Specifically, they imply that among formal sector producers, low plant or job turnover, high mark-ups, and the frequent survival of inefficient plants are symptoms of high sunk entry costs. Similarly, Hopenhayn and Rogerson’s (1993) simulations imply that persistence in

---

market shares and low turnover among formal sector firms are symptoms of binding severance laws or restrictions on the use of temporary workers.

In addition to studies of the size distribution of firms, at least three empirical literatures provide evidence on these symptoms. The first summarizes the extent of productivity dispersion, usually in the context of efficiency frontier estimation. The second, relatively recent literature documents the extent of plant turnover, and in some cases relates this turnover to productivity growth. Finally, an older literature on industrial concentration is potentially relevant. Let us take each in turn.

4.2 Is Productivity Dispersion Higher in LDCs?

Many analysts have studied the amount of productivity dispersion in LDCs. A sampling of results is presented in Tables 2 and 3, and compared to those from a recent multi-country study of the OECD. Each study is done

---

**Table 3**

**Stochastic Frontiers: Average Technical Efficiency by Industry**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>0.67</td>
<td>0.46</td>
<td>(weaving only) min: 0.618 max: 0.766</td>
<td>0.642</td>
</tr>
<tr>
<td>Textiles and garments</td>
<td>0.626</td>
<td></td>
<td>(apparel)</td>
<td>0.642</td>
</tr>
<tr>
<td>Footwear</td>
<td>0.42</td>
<td></td>
<td></td>
<td>0.554</td>
</tr>
<tr>
<td>Wood and furniture</td>
<td>0.42</td>
<td></td>
<td></td>
<td>0.984</td>
</tr>
<tr>
<td>Metal products</td>
<td>0.719</td>
<td>0.51</td>
<td></td>
<td>0.987</td>
</tr>
<tr>
<td>Machine tools</td>
<td>0.727</td>
<td>0.704</td>
<td>(agricult. hand tools only)</td>
<td></td>
</tr>
<tr>
<td>Plastic products</td>
<td>0.820</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor vehicles</td>
<td>0.846</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* All figures are estimates of $E(e^{-u})$, where the inefficiency measure $u$ is assumed to follow a half-normal distribution.

** Differences in estimated average efficiency reflect differences in the way that labor is measured, and whether plant characteristics like size and foreign ownership dummies are included in the production function.
by estimating the “frontier” production technology, which defines the maximum amount of output, \( y^* \), attainable from a given vector of inputs, \( x \): \( y^* = f(x) \). Then, for observed combinations of output and inputs at the \( i^{th} \) plant \( (y_i, x_i) \), the ratio \( y_i / f(x_i) \) is interpreted either as an efficiency index itself, or as an efficiency index contaminated by measurement error and transitory shocks beyond the control of plant managers. These two approaches are known as the “deterministic frontier” and the “stochastic frontier” approach, respectively. Cross-plant average efficiency levels and standard deviations in efficiency levels are the most commonly reported summary measures of an industry’s performance. These bear a negative monotonic relationship to one another in most cases, so I report only the former.

27 The literature can be further sub-divided according to whether \( f(x_i) \) is estimated parametrically or non-parametrically (known as “data envelopment analysis”), and whether econometric or programming techniques are used. William Greene (1993) provides a recent summary of the various approaches to efficiency measurement.

<table>
<thead>
<tr>
<th>Caves et al. (1992)</th>
<th>Corbo and de Melo (1986), Chilean firms</th>
<th>Kalirjian and Tse (1989), Malaysian firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food processing</td>
<td>0.713</td>
<td>0.73</td>
</tr>
<tr>
<td>Shoes</td>
<td>wrong skewness</td>
<td></td>
</tr>
<tr>
<td>Printing</td>
<td>wrong skewness</td>
<td></td>
</tr>
<tr>
<td>Soap</td>
<td>0.627</td>
<td></td>
</tr>
<tr>
<td>Machine tools</td>
<td>wrong skewness</td>
<td></td>
</tr>
<tr>
<td>Agricultural machinery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plastic products</td>
<td>wrong skewness</td>
<td></td>
</tr>
<tr>
<td>Motor vehicles</td>
<td>wrong skewness</td>
<td></td>
</tr>
<tr>
<td>Textiles: spinning</td>
<td>wrong skewness</td>
<td></td>
</tr>
<tr>
<td>Apparel and weaving</td>
<td>0.649</td>
<td></td>
</tr>
<tr>
<td>Saw mills</td>
<td>0.652</td>
<td></td>
</tr>
<tr>
<td>Furniture</td>
<td>wrong skewness</td>
<td></td>
</tr>
<tr>
<td>Japan, cross-industry average of 144 industries</td>
<td>0.699</td>
<td></td>
</tr>
<tr>
<td>Korea, cross-industry average of 128 industries</td>
<td>0.672</td>
<td></td>
</tr>
<tr>
<td>UK, cross-industry average of 72 industries</td>
<td>0.680</td>
<td></td>
</tr>
<tr>
<td>Australia, cross-industry average of 91 industries</td>
<td>0.699</td>
<td></td>
</tr>
<tr>
<td>US, cross-industry average of 67 4-digit industries</td>
<td>0.671</td>
<td></td>
</tr>
</tbody>
</table>
Some caveats are in order. First, these studies are done at differing levels of aggregation. One would expect that the finer the industry, the less dispersion due to pooling heterogeneous technologies. Second, there are differing degrees of measurement error in outputs and inputs. Much of the cross-plant heterogeneity in capital and labor is unobserved by the econometrician, and most studies describe output in terms of revenue rather than physical product, blurring the distinction between factor productivity and price-cost mark-ups. Third, as is well known, the results depend to a large degree upon whether stochastic or deterministic frontiers are used, and upon the assumed distribution of the error terms (Vittorio Corbo and Jaime de Melo 1986).

Finally, unless they are estimated with panel data, stochastic frontier models separate technical inefficiency from noise by treating \( \ln[\frac{y_i}{f(x_i)}] \) as the sum of two orthogonal error components—one reflecting inefficiency and the other reflecting measurement error or shocks beyond the control of managers. Typically, the negative of the inefficiency component (hereafter denoted \( u \)) is assumed to have a half-normal, gamma or exponential distribution, and the noise component is assumed to have a normal distribution. Greater skewness—measured by the negative of the third moment of the compound error—thus implies more productivity dispersion. However, the data often imply that the distribution of \( \ln[\frac{y_i}{f(x_i)}] \) is skewed in a way that is inconsistent with these assumptions, so in practice many industries do not fit the model. Such industries are typically dropped from the analysis, and the reported average efficiency levels are based only on the industries that remain.

To control for differences in methodology, I have sorted studies according to whether they presume deterministic or stochastic frontiers, and wherever possible, in the latter case I have used the results based upon the half-normal distribution for the efficiency component of the error term. Among the deterministic frontier studies, there is still some variation in the methodologies across studies because some use linear programming to identify the production function while others use quadratic programming. More importantly, some (like Page 1980) impose a distribution on the efficiency measures, while others do not.

The deterministic frontier studies (Table 2) generally yield lower average efficiency levels than the stochastic frontier studies (Table 3), since the former attribute all unexplained variation in \( y \) to inefficiency. Unfortunately, they are also very sensitive to the specific assumptions behind the calculations, and do not appear to convey any clear messages. Notice, for example, that Corbo and de Melo’s (1986) deterministic frontier estimates imply that Chile was very inefficient relative to other countries, but their stochastic frontier estimates imply Chile was average. Given this sensitivity, as well as the lack of good comparator studies from industrialized countries, I shall hereafter focus on the stochastic frontier results.

Doing so reveals a surprising pattern. It is often observed that the cross-firm variance in productivity levels is high in developing countries—e.g., Pack (1988), Evenson and Westphal (1995), Magnus Blomstrom and Ari Kokko (1997). Nonetheless, Table 3 suggests that average deviations from the efficient frontier are not typically larger than what we observe in the high-income countries studied by Caves et al. (1992). The standard methodology, when it “works,” yields mean technical efficiency levels around 60 to 70 percent of the best.
practice frontier in both regions. Hence the studies surveyed provide little support for the view that LDC markets are relatively tolerant of inefficient firms.  

One exception is provided by Biggs, Shah, and Srivastava (1995), who report an unusually large amount of productivity dispersion in Ghana, Zimbabwe, and Kenya. However, these results are based on relatively broadly-defined industries, so they may be a simple consequence of aggregation bias. In a particularly detailed study, Pack (1984) finds average deviations among Kenyan textiles producers comparable to those in other studies, even though his methodology is based on deterministic frontiers. Similarly, Page (1980) finds dispersion levels typical of other countries in his early study of Ghana.

Although the studies summarized in Table 3 do not find higher productivity dispersion in LDCs, they are not very informative. Most of them are based on outdated methodologies. With a few exceptions they rely on cross-sectional data, and hence must infer efficiency dispersion from the skewness of the production function residuals. Further, because most measure output as real revenue, they misattribute cross-plant mark-up differences to productivity dispersion. Finally, for lack of data, they typically equate high productivity with superior performance, ignoring many of the costs that firms incur to enhance their technical efficiency. It is surely not optimal to continually be the most productive firm (George Stigler 1976). But without measuring the net present value of firms’ efficiency-enhancing expenditures and the associated changes in their productivity trajectories, one cannot know whether observed patterns of productivity dispersion are problematic.

4.3 Plant and Job Turnover in LDCs

The literature on plant and job turnover may be a better place to look for evidence on the strength of competitive pressures in the LDCs. If extensive regulation and taxation combine with credit market problems to keep small firms from challenging their entrenched larger competitors, we should observe few firms graduating from informal to formal status. Further, those firms that do graduate should show relatively little mobility up the size distribution, and markets shares should be relatively stable among the largest firms.

Unfortunately, turnover figures are unavailable for many of the countries where one would expect the policy regime to inhibit flux. However, a handful of studies on Latin American, East Asian and North African countries provide some preliminary evidence. These studies typically document entry rates, exit rates, net job creation and net job destruction patterns among the population of plants with at least ten workers. (Variable definitions are provided in the footnote to Table 4.) Most firms above the ten-worker threshold participate wholly or partly in the formal sector, so measured entry rates crudely describe formal sector entry—both through new plant creation and through the graduation of informal plants to formal status. Similarly, the job turnover rates reported give us
a crude sense of the stability of market shares among formal sector firms.

Surprisingly, although Cox-Edwards (1993, p. ii) argues that Latin American countries “have a long tradition of trying to protect employment stability,” there appears to be more plant and job turnover in these developing countries than others have found in the United States and Canada (Table 4).31 Over a five-year interval, entering plants with at least ten workers captured 15 percent of the market in Chile, and entering

\[ * \]

\[ \text{Let } N_t \text{ be the number of plants observed in year } t; E_t \text{ be the number of plants observed in year } t \text{ but not } t-1; \text{ and } X_t \text{ be the number of plants observed in year } t-1 \text{ but not in year } t. \text{ Then the entry rate is } E_t/N_t-1 \text{ and the exit rate is } X_t/N_t-1. \text{ The plant turnover rate is the average of these two statistics. Similarly, the rate of gross job creation is the number of jobs at entering plants plus the number of new jobs at expanding plants, divided by initial number of jobs, and the gross job destruction rate is the number of jobs that disappear as plants contract or exit divided by the initial number of jobs. The sum of these two rates is the job turnover rate.} \]

\[ \text{Sources:} \]

\[ ^a \text{ Roberts and Tybout (1996), Tables 2.3, 9.3, 9.5, 10.3, 10.4, and 12.3.} \]

\[ ^b \text{ Sukkyan Chung (1999), Tables 3.1 and 3.2.} \]

\[ ^c \text{ Xiaomin Chen (1997), Tables 2.7 and 2.8; Aw, Chen, and Roberts (1997), Table 1.} \]

\[ ^d \text{ Timothy Dunne, Roberts, and Larry Samuelson (1988), Tables 2, 3, 4, 8, and 9. Figures are average rates of new entry rates across 4-digit industries.} \]

\[ ^e \text{ John Baldwin, Dunne and John Haltiwanger (1998), Table 1. These figures are for 1973–92.} \]

\[ ^f \text{ Dunne, Roberts and Samuelson (1989), Table 1.} \]

\[ ^g \text{ The Taiwanese data set describes firms rather than plants.} \]

31 Among other distinctive features, Cox-Edwards (1993) notes that “The Latin American legislation, with a few exceptions, including Mexico, is very strict in limiting the use of temporary contracts . . . firms cannot rely on a mix of permanent and temporary labor force, as is the case in Japan and increasingly the United States . . .” (p. 14). Hence it is difficult to avoid severance payments by relying on temporary workers. On the other hand, as Cox-Edwards emphasizes, severance payments are often legally tied to number of years on the job, so, subject to the temporary worker constraints, firms may be encouraged to “maintain a very young work force with high rotation . . .” (p. iii)
plants with at least ten workers captured 20 percent of the market in Colombia. On the other hand, in the more inclusive population of plants with at least five workers, entrants in the United States captured an average of only 10 percent of the market. In terms of job creation and job destruction, Chile and Colombia average 27 and 25 percent annual turnover rates, respectively, while the United States and Canada average 19 percent and 22 percent, respectively.

Outside Latin America, some analysts have found even more flux in plants and jobs. In Morocco the annual manufacturing job turnover rate was 31 percent. In Korea and Taiwan, over five-year intervals, new entrants captured an average of 33 percent and 44 percent of the market, respectively, compared to 10 percent in the United States. Finally, although studies of the complete size distribution are unavailable for sub-Saharan Africa, Liedholm and Mead (1995) find that turnover rates among micro and small enterprises are very high, ranging from 19 to 25 percent per annum (not in Table 4).

Why is there so much flux in these LDCs? In some cases—especially in Latin America—high turnover partly reflects the relatively dramatic business cycles found there. In others—especially Korea and Taiwan—it partly reflects rapid expansion of the manufacturing sector. But even if one focuses on the minimum of the entry rate and the exit rate, turnover is relatively rapid in the developing countries that have been studied.

Another part of the explanation lies with Engel effects and low levels of human capital, which encourage turnover by skewing the output mix toward simple products with relatively low start-up costs, like baked foods, footwear, apparel, and metal products (Figure 2). The dominance of these sectors and technologies is probably amplified by macro uncertainty, which creates incentives to be flexible in terms of productive capacity.

Finally, policies seem to matter. Market share turnover rates are higher in Korea and Taiwan than they are in Latin America, where labor markets are relatively regulated. Indeed, although ambiguities remain, turnover rates appear to be highest in Taiwan. Taiwan’s labor markets are least regulated among the sample countries (Joseph Lee and Young-Bum Park 1995), and sunk entry costs are relatively modest there because the business environment makes subcontracting easy (Brian Levy 1990).

Of course, the turnover rates discussed above don’t reveal which plants in the population are expanding or contracting. It is possible that nearly all of the turbulence takes place among plants in the 10–50 worker range, and that these moderately small producers never seriously challenge the larger, entrenched incumbents. Indeed, one might argue that high turnover rates in LDCs simply reflect the relative importance of small and medium enterprises there (Table 1), and need not imply that large firms’ market shares are more at risk.

32 Since turnover takes place mainly among small plants, the contrast between these North and South American countries would have been even greater if all studies had been done on a comparable basis. Turnover figures from European countries are available, but they are reported with insufficient documentation to be useful for present purposes (John Cable and Joachim Schwalbach 1991).

33 The market share figures for Korea cover all plants with at least five workers, as are the United States figures. Note, however, that the Taiwanese data describe all firms.

34 Comparisons of Taiwan with other countries are unfortunately clouded by differences in sample coverage (see previous footnote). Further evidence that Taiwan has relatively high turnover will be introduced shortly.
For some of the LDCs that have been studied we cannot rule out this possibility. For example, Chilean and Colombian plants with at least ten workers lost 15 and 20 percent of their markets, respectively, over a five-year period (Table 4). But plants with less than 50 workers account for more than half of total manufacturing employment in these countries (Table 1), so all of the market share loss could have come at the expense of small producers.35 This is not true in Korea, where entrants captured 32.5 percent of the output market and accounted for 46.5 percent of employment after five years, while plants with less than fifty workers accounted for only 29.9 percent of manufacturing employment. It is even less true in Taiwan. There, plants observed in 1981 had lost 44 percent of the market to new entrants by 1986 and they had lost 63 percent of the market by 1991, but plants with less than one hundred workers accounted for only 49 percent of employment.

4.4 Turnover and Productivity Growth

High turnover does not necessarily imply that inefficient producers are rapidly driven from the market. For example, when the Argentine exchange rate regime collapsed in the early 1980s, it left many firms with dollar-denominated debt in serious trouble, and the resulting exit patterns had little to do with productive efficiency (Eric Swanson and Tybout 1988). It is therefore worth inquiring how well turnover “cleanses” LDCs of their least productive plants.

Several studies have quantified the effects of exit and entry on sector-wide productivity growth. Each is subject to the measurement-error problems mentioned in Sections 3.2 and 4.2, but the findings seem plausible and are generally consistent with one another. In Chile and Colombia, as in developed countries, measured productivity among exiting plants is much lower than it is among incumbents (Lili Liu and Tybout 1996; Liu 1993; and Tybout 1992b). Indeed, the productivity of exiting Chilean plants has often begun to deteriorate several years before they actually exit (Liu 1993)—a phenomenon that Zvi Griliches and Haim Regev (1995) dubbed the “shadow of death” effect in their study of Israeli turnover. Similarly, Taiwanese plants doomed to disappear in the next five years exhibit below-average efficiency (Aw, Xiaomin Chen, and Mark Roberts 1997). So there is evidence that a shakedown process is at work.

However, in Chile and Colombia, entering plants are also less productive than incumbents on average. Further, neither entrants nor dying plants account for more than 5 percent of total output in a typical year.36 So inefficient plants are being replaced with plants that are only slightly more efficient, and neither group is a source of much production. This implies that if the turnover process were suddenly arrested, the impact on productivity would initially be small.

Nonetheless, over time the costs of policies that prevent turnover quickly mount for several reasons. First, the “shadow of death” effect suggests that exiting plants are on a downward trajectory, and might well continue to get worse. Second, entering cohorts typically undergo a shakedown period in

35 For lack of information, I am ignoring the distinction between employment shares and output shares here. Among small plants the former typically exceed the latter.

36 The low average productivity of entering plants might seem at odds with the results I discussed earlier which suggested small plants are not much less efficient than large ones. These two findings are not contradictory because, while most new plants are small, most small plants are not new.
which the least efficient entrants drop out and the survivors quickly improve their productivity. Liu and Tybout (1996) find that this process brings the average productivity of new cohorts up to industry-wide norms after three or four years in Colombia, and Aw, Chen, and Roberts (1997) find similar catch-up patterns in Taiwan, although the process is not complete there after five years in some industries. Finally, while the firms turning over account for a small share of production in any one year, the cumulative effects of turnover on the population of plants quickly mount.

To empirically link the business environment with turnover-based productivity gains, and to be rigorous about it, one would need to fit a dynamic structural model of industrial evolution to firm-level panel data. That has yet to be accomplished, so in the meantime one is tempted to look at cross-country correlations between productivity gains and policies that affect turnover. Unfortunately, even this is problematic because only two country studies have calculated turnover-based productivity figures on a roughly comparable basis. We are thus left with the single tantalizing observation that turnover-based productivity growth has been higher in Taiwan (3.2 percent over five years) than in Colombia (2.2 percent over five years). Those who are predisposed to do so might conclude that Taiwan’s relatively laissez-faire policies are responsible.

4.5 Concentration, Price-Cost Margins, and Market Power

Industrial product markets are relatively concentrated in developing countries, and from this fact some observers have inferred—contrary to the turnover studies discussed above—that LDCs lack competition. One rationale for this inference is that high concentration results from high entry costs and institutional constraints on labor markets, which limit the number of players and insulate them from competitive pressures. Another rationale is that high concentration in developing countries traces to small domestic markets rather than restrictive policies, but it nonetheless increases the sustainability of collusive arrangements.

Neither justification for inferring market power from concentration seems compelling. The first rationale does not fit those developing countries where evidence on turnover is available. High sunk entry costs and labor market constraints should depress turnover at the same time that they increase concentration, but we have seen that these countries show rapid flux in plants and jobs. The second rationale is correct insofar as market size is a powerful predictor of concentration. Indeed, in Latin America two-thirds of the cross-country variation in industrial concentration measures is explained by the logarithm of GDP alone. But concentrated markets

37 These figures are weighted averages of industry-specific results reported in Aw, Chen, and Roberts (1997) and Liu and Tybout (1996), respectively.

38 Norman Lee (1992) surveys the empirical literature on concentration in LDCs so I will not repeat the exercise here. Theorists have also been known to view high concentration in the LDCs as signaling relatively uncompetitive markets there (Paul Krugman 1989, Rodrik 1988).

39 Simulations of an industrial evolution model with noncompetitive market structures verify that when countries differ because of sunk entry costs, differences in concentration are positively correlated with market power and monopoly rents (Pakes and Paul McGuire 1994).

40 This is the $r^2$ I obtain using Patricio Mellor’s (1978) concentration measures, which were constructed the same way for ten Latin American countries using their industrial census data. A number of studies have commented on the negative correlation between market size and concentration—Lee’s (1992) survey provides further details.
in the presence of vigorous turnover need not be conducive to collusive behavior.\textsuperscript{41}

If we eschew concentration as a signal of market power, what other evidence is available? Some studies have exploited Richard Schmalensee’s (1985) methodology, which “. . . amounts to asking whether cross-plant variations [in price-cost margins] are due to industry-wide effects or to plant-specific market shares. Efficient plants should be larger and have higher profits, so a positive correlation is generally expected between market shares and price-cost margins, regardless of whether firms have market power.” (Roberts and Tybout 1996, p. 196) On the other hand, persistent cross-\textit{industry} variation in mark-ups, controlling for capital intensity, suggests that profit differentials are not arbitraged away; that is, that entry costs in some sectors limit competition. There are many problems with the studies that apply Schmalensee’s logic to LDCs, including the poor correspondence between their profit measures—price-cost margins—and economic profits (Franklin Fisher and John McGowan 1983). But taken at face value, the results for Chile, Colombia, and Morocco show no more evidence of market power than Schmalensee (1985) found in the United States (Roberts and Tybout 1996).\textsuperscript{42}

4.6 \textit{The Bottom Line}

To summarize, because of institutional entry barriers, labor market regulations, poorly functioning financial markets and limited domestic demand, the industrial sectors of developing countries are often described as insulated, inefficient oligopolies. To date, however, there is little empirical support for this characterization. Turnover is substantial in the countries that have been studied, unexploited scale economies are modest, and evidence of widespread monopoly rents is lacking.

The above notwithstanding, it would be foolish to conclude that market power is a nonissue in developing countries. Turnover studies and cross-plant studies of profitability give one a general sense of the extent of competition, but they cover a limited and perhaps unrepresentative set of countries. Further, they are unlikely to detect isolated pockets of noncompetitive behavior. For example, in Chile and Colombia during the 1970s, a handful of closely held conglomerates controlled large shares of certain industries, as well as portions of the financial sector (Fernando Dahse 1979; Superintendencia de Sociedades 1978). More recently, cozy relationships between such conglomerates and the state have attracted attention in East Asia (Ashoka Mody 1998). Finally, many developing countries have privatized natural monopolies during the past decade, and where efficient regulatory agencies have not sprung up to oversee them, noncompetitive practices may be on the rise. Careful case studies that collect detailed price data and monitor the behavior of the individual players are probably the only means through which convincing conclusions about these problems can be reached.

5. \textit{Trade Protection, Market Structure, and Productivity}

Even in those countries where competition is vigorous, it is nonetheless imperfect. Entry and exit costs matter, and
products are differentiated. Further, learning spillovers and other externalities are surely present in some form. Hence protectionist trade policies, where they still exist, may do more than affect domestic relative prices and intersectoral resource allocations. They may change *intra*-industry mark-ups, productivity, or productivity growth. In this section, after briefly recounting the relevant theoretical literature, I consider the firm-level econometric evidence from LDCs on each possible effect.

5.1 Possible Effects of Trade Policy

Static arguments: There are numerous static arguments why trade protection might affect the performance of domestic firms in LDCs. Most involve the effects of trade policy on the competitive pressures that these firms face, the size of the market that they operate in, or both. Firms’ responses often depend upon whether entry and exit barriers are substantial, whether scale economies—internal or external—are important, and whether protection takes the form of tariffs or quantitative restrictions.43

I will limit myself to several examples. Consider a tradeable goods industry with substantial entry barriers, composed of Cournot-competing firms. If the industry has zero net exports under free trade, the main effect of import prohibitions is to eliminate the threat of foreign competition. Domestic firms may exploit their enhanced market power by curtailing production and increasing their price-cost mark-ups, perhaps sacrificing some scale efficiency in the process.

On the other hand, if the industry begins from substantial import penetration, the dominant effect of protection may be to increase the market size for domestic producers. Firms are likely to respond by expanding, perhaps exploiting scale economies as they do so. (Mark-ups may still rise.) In either scenario, the higher profits that result from protection may allow relatively inefficient firms to survive, driving up productivity dispersion. Alternatively, if we drop the assumption of prohibitive entry barriers—which seems sensible, given our findings in Section 4—the higher profits may eventually entice new, inefficiently small domestic producers to enter (Krugman 1979).44

External economies of scale further expand the list of possible effects of trade policy on productivity. Suppose, for example, that the external economies occur at the industry level, and are national rather than global.45 Then the net effect of trade liberalization depends upon which sectors expand and which contract, as well as the magnitude of traditional gains from comparative advantage effects. It is possible that the losses can outweigh the gains (e.g., Helpman and Krugman 1985, ch. 3).

Finally, when employee effort is a choice variable, trade policy can affect the amount of “managerial slack” or “X-inefficiency” among manufacturers. The dominant view among development economists is that protection induces managers in import-competing industries

---

43 Many of the relevant models are summarized in Elhanan Helpman and Krugman (1985) and a number of the seminal contributions are collected in Gene Grossman (1992).

44 Keith Head and John Reis (1999) summarize the analytical and empirical literature on trade liberalization and the size distribution of firms, while providing some new evidence from Canada.

45 Industrial expansion generally deepens the market for specialized labor, material inputs, and networked support services. Thus, even if no technology spillovers take place, external scale economies at the industry level may be present when there are increasing returns to scale in the production of these inputs, or when risk-averse specialized workers prefer regions with many job opportunities (e.g., Francisco Rivera-Batiz and Luis Rivera-Batiz 1990; Francis Stewart and Ejaz Ghani 1992).
to relax and enjoy the “quiet life.” In early versions of the argument, protection increases profits among domestic firms. This relaxes the consumption-leisure budget constraint faced by their managers, who respond by choosing more of both if they are on the backward-bending portion of their labor supply schedule (W. Max Corden 1974). In more recent treatments, protection affects the payment schedule that owners (principals) must offer to managers (agents) to induce them to reveal their endowed abilities.\footnote{Given his or her endowed ability, each manager chooses an effort level in response to the reward structure and market conditions. By the revelation principle, contracts that induce managers to be truthful about their (unobservable) abilities yield at least as high a value to the owner as any other mechanism.} If the cost to owners of truthful revelation rises with protection, they may opt for equilibria at lower output and effort levels, but the effect of protection on effort is sensitive to modeling details (e.g., Neil Vousden and Neil Campbell 1994).

**Dynamic arguments:** Further effects of trade policy on performance have been demonstrated in explicitly dynamic frameworks. Again, most anything can happen, depending upon modeling assumptions and the particular policy experiment. One issue that has attracted attention is whether trade protection will induce technologically backward producers to invest in catching up. In theory it may, if it increases the effective market size and the associated pay-off from marginal cost reductions for domestic firms (Kaz Miyagiwa and Yuka Ohno 1995; Rodrik 1992). On the other hand, protection may facilitate collusion among domestic producers and induce them to collectively stick with backward technologies (Rodrik 1992). A modest permanent quota may also delay technology adoption because, with continuously binding quantity constraints, foreign suppliers do not cut back their shipments to the domestic market when the home firm becomes more efficient (Miyagiwa and Ohno 1995). Catch-up models describe a one-time transition from dated to new technologies, but they do not link trade policies to ongoing productivity growth. For that, theorists have developed general equilibrium frameworks with continual knowledge production and diffusion. In such models, protection changes the relative prices of the inputs involved in product development, affects the set of imported products that innovators compete with, and affects the ease with which domestic innovators can access foreign technical expertise.

Whether protection reduces ongoing productivity growth in these models depends partly upon the way in which knowledge diffuses.\footnote{Another key issue is the strength of spillover effects. Diffusion of knowledge through technology purchases or licensing agreements is not enough to establish a link between steady state growth and trade policies. Knowledge spillovers are typically needed, and they must be strong enough that the private return to innovation does not fall with increases in the stock of knowledge (Charles Jones 1995).} Suppose that trade policy does not affect the ease with which foreign knowledge can be accessed, perhaps because it is readily available over the internet. Further, suppose that to efficiently deploy technical knowledge in LDCs, there is no substitute for learning-by-doing in the high-tech sectors and the spillovers it generates. Then trade protection may improve productivity growth and welfare if it promotes the high-tech activities that generate the highest learning rates and the most valuable spillovers.\footnote{Although they have not endorsed it, this argument for protection has been formalized by Krugman (1987), Nancy Stokey (1988), Alywn Young (1991) and Grossman and Helpman (1991).}

On the other hand, if domestic producers acquire some of their knowledge...
through exposure to foreign clients, 
technologically sophisticated imports, 
or knowledgeable competitors, protec-
tion may slow growth by constricting 
important channels of knowledge trans-
misison (e.g., Grossman and Helpman 
1991, ch. 6). Similar comments apply to 
policies that discourage foreign direct 
investment if the local presence of mul-
tinational plants facilitates technology 
diffusion.

Overall, the most striking conclusion 
that emerges from the analytical litera-
ture discussed above is that almost any-
thing can happen when a country pro-
tects its manufacturers, depending upon 
the assumptions one invokes. Hence 
many empiricists have attempted to de-
termine what happens in practice by 
studying patterns of association be-
tween trade policy, pricing behavior, 
productivity, and productivity growth. 
Others have attempted to chip away at 
ambiguities by asking which modeling 
assumptions best describe the data. Let 
us now consider the evidence.

5.2 Openness and Pricing Behavior: 
The Evidence

One of the more robust analytical re-
sults on trade with imperfect competi-
tion is that policies that constrain im-
ports tend to increase the market power 
of domestic producers. To look for 
evidence of this phenomenon, many re-
searchers have regressed price-cost 
margins on proxies for import competi-
tion or trade protection, usually looking 
across industries at a point in time.49 
The correlation between import pen-
etration (trade protection) and margins is 
typically negative (positive), and the 
standard interpretation is that foreign 
competition squeezes monopoly rents, 
or “disciplines” the pricing behavior of 
domestic producers.

In a variant on this theme, a number 
of authors have recently used Robert 
Hall’s (1988) methodology for measur-
ing mark-ups to gauge the effects of im-
port competition on pricing behavior.50 
The typical exercise is to regress output 
growth on a share-weighted average of 
input growth rates, and interpret the 
coefficient on input growth as a mono-
tonic function of the price-cost mark-
up. Allowing the coefficient to shift 
with trade liberalization, most studies in 
this genre find that openness is associ-
ated with reductions in price-cost mar-
gins, and they interpret this to support 
the “import discipline” hypothesis.51

However, even if one ignores the 
econometric problems, other interpre-
tations are possible. Suppose that in-
creased import-penetration reflects real 
exchange rate appreciation, which 
squeezes output prices relative to input 
prices in the tradeable goods industries. 
In the short run, so long as revenues 
still cover variable costs, all firms— 
competitive or otherwise—will produce 
at lower margins and the competitive 
sectors will make negative economic 
profits. Alternatively, suppose that 
heightened import penetration reflects 
the removal of trade barriers. Then 
relative prices have been twisted in fa-
vor of exportables and against import-
able. Stolper-Samuelson effects should 
drive down the relative price of the in-
put used intensively by the latter sec-
tors, which is likely to be capital in the 
developing countries. Similarly, in 
cross-sectional studies, suppose that 

49 Norman Lee (1992) surveys this literature. 
50 Levinsohn (1993), Harrison (1994), Faezeh 
Foroutan (1996) and Pravin Krishna and 
Devashish Mitra (1997) study Turkey, Cote 
d’Ivoire, Turkey (again), and India, respectively. 
51 This methodology is likely to suffer from si-
multaneity bias because transitory productivity 
shocks appear in the disturbance term, and are 
likely to be correlated with input growth. Appropriate 
instrumnts are nearly impossible to find 
(Thomas Abbott, Zvi Griliches, and Jerry Haus-
man 1989).
protection is attracted to the industries in the country’s comparative disadvantage, which are likely to be capital intensive. These same industries should exhibit relatively high price-cost margins simply because relatively large amounts of capital are used per unit output.

5.3 Openness and Productivity Levels: The Evidence

In addition to falling mark-ups, empiricists tend to find that trade liberalization is associated with rising average efficiency levels (Mieko Nishimizu and Page 1982; Tybout, de Melo, and Corbo 1991; Tybout and Westbrook 1995; Harrison 1996). Similarly, protected industries tend to exhibit heightened productivity dispersion (Tybout, de Melo, and Corbo 1991; Haddad 1993; Haddad and Harrison 1993). The standard interpretation of these results is that foreign competition drives inefficient domestic producers to exploit scale economies, eliminate waste, adopt best practice technologies, or shut down.

However, as with the mark-up studies, a number of caveats apply. First, simultaneity bias creates the usual problems. Inefficient, influential firms often lobby for protection, and sometimes they succeed. Further, as already discussed, in most firm-level data sets output is measured as revenue divided by an industry-wide deflator. So reductions in measured “productivity” dispersion may simply mean that mark-ups have fallen the most among firms with the largest initial margins. That is, when efficiency is equated with low dispersion—as it is in the efficiency frontier literature—improvements in productivity cannot be distinguished from the mark-up squeeze often associated with trade liberalization. Conversely, if trade liberalization is associated with a major devaluation, the favorable twist in prices for tradeables should increase their profitability, at least in the short run. This looks just like an increase in average efficiency among tradeable goods producers if physical units of output cannot be observed.

Going beyond the association between measured efficiency and openness, a number of authors have attempted to determine why the two are correlated. Their findings suggest that internal scale effects are not the main reason. If trade liberalization forces inefficiently small firms down their cost curves, one should observe plant sizes rising in import-competing sectors as protection is removed. (This mechanism is built into the simulation models referenced in footnote 16.) However, micro panel studies consistently find that increases in import penetration are associated with reductions in plant size, as are reductions in protection (Mark Dutz 1996; Roberts and Tybout 1991; Tybout and Westbrook 1995). Thus liberalization may work against scale efficiency, at least in the short run. Nonetheless, the impact on efficiency of these plant size adjustments is probably small, since adjustments take place mainly among large plants that are operating in the constant returns range of their cost curves (Tybout and Westbrook 1995).

Although external scale economies are probably present, they too are unlikely to account for large protection-related efficiency effects. Using plant-level panel data and the methodology developed by Ricardo Caballero and Richard Lyons (1990), Cornell J. Krizan (1997) finds significant external returns to scale in many Moroccan industries. Other evidence that external scale effects are present comes from the locational choices of new firms (e.g., Vernon Henderson and Ari Kuncoro 1996).

52 Other evidence that external scale effects are present comes from the locational choices of new firms (e.g., Vernon Henderson and Ari Kuncoro 1996).
computable general equilibrium model of Morocco developed by Thomas Rutherford, E. E. Rustrom, and David Tarr (1993) and simulates the effects of trade liberalization vis-à-vis Europe. He finds that external economies compound the gains from liberalizing, but the effects are quite small. Further, to obtain large efficiency gains or losses one would have to assume implausibly large external returns.

In sum, when trade liberalization improves productive efficiency, it is probably largely due to intra-plant improvements that are unrelated to internal or external scale economies. The elimination of waste, reductions in managerial slack, heightened incentives for technological catch-up, and access to better intermediate and capital goods are all possible explanations, but there is little direct evidence on the importance of any of these. Detailed analysis of task-level efficiency and technological choice within narrowly defined industries—before and after a major change in trade policy—is probably the most promising direction for further work on the topic.54

5.4 Openness and Productivity Growth: The Evidence

Static and dynamic effects of trade policy are conceptually distinct in that the latter involve a time dimension. However, all responses to policy take time, even those that can be analytically described with a static model. Given the short time periods spanned by micro data, it is thus rarely possible to distinguish transitory one-shot adjustments in productivity levels from lasting changes in the rate of productivity growth. Hence, to assess the relevance of the dynamic analytical models I mentioned in section 5.1, I will limit my discussion to the issue of how technology diffuses.

Technology transfers through trade. As already noted, outward-oriented policies are more likely to facilitate long-run growth if technology diffuses through international transactions. For example, LDCs may acquire new technologies by de-engineering imports, or simply by deploying the innovative intermediate and capital goods that they acquire in foreign markets. They may also learn about product design and new technologies or management techniques from the foreign buyers to whom they export. Once acquired through these channels, new foreign technologies may diffuse to other domestic firms not directly engaged in trade. Are these processes empirically important?

There is very little micro-econometric evidence on the productivity enhancing effects of importing sophisticated intermediate and capital goods, although the fact that many LDCs import most of their machinery and equipment speaks for itself. Several studies do report a positive correlation between access to imported intermediate goods and performance (Heba Handoussa, Nishimizu, and Page 1986; Tybout and Westbrook 1995), and Robert Feenstra, James Markusen, and William Zeile (1992) report evidence that Korean firms improved their productivity by diversifying their input bundles. Thus imported capital and intermediate goods may be an important channel through which trade diffuses technology, but further work is clearly needed to quantify the effects.

More detailed evidence is available on technology acquisition through exporting.

---

53 In the scenario with the largest externality effects, the positive effects of trade liberalization on welfare increase from a 0.9 percent gain to a 1.1 percent gain.

54 Pack (1984) and Mody et al. (1991) provide excellent examples of research at this level of detail, but neither study directly examines the link between performance and trade reforms.
In support of a “learning by exporting” effect, most studies that compare the productivity of LDC exporters with that of others in the same industry and country find that exporters do better. Case studies confirm that OECD buyers sometimes provide their LDC suppliers with blueprints and technical advice (e.g., Yung Rhee, Bruce Ross-Larson, and Garry Pursell 1984). But firms that are relatively efficient—perhaps for the reasons described in the industrial evolution literature—are also relatively likely to self-select into foreign markets (Sofronis Clerides, Saul Lach, and Tybout 1998). Hence the cross-sectional correlation between exporting and efficiency may reflect causality in either direction, or both.

Several recent studies attempt to address the causality issue by tracking firms through time and asking, first, whether those that became exporters were more efficient beforehand, and second, whether exporters showed improvement relative to industry norms after entering foreign markets. Most find that exporters were substantially more efficient than non-exporters before they started selling abroad, so the higher efficiency of exporters appears to be at least partly a self-selection effect. These studies also find that in most industries the efficiency gap between exporters and non-exporters does not grow over time, suggesting that learning is not a general phenomenon. However, firms in several industries do exhibit relative efficiency gains after becoming exporters, so the learning-by-exporting hypothesis cannot be ruled out entirely.

Regardless of whether they acquire their expertise from abroad, technology may also diffuse from exporters to non-exporters in the same country, region or industry through demonstration effects, skilled worker training (and subsequent labor turnover), or expertise imparted to their local suppliers. Looking at the intensity of exporting activity through time, Clerides, Lach, and Tybout (1998) find that when many firms have been exporting from a particular region, all firms in that region tend to enjoy lower average costs. Spillovers are one interpretation, but this finding may simply reflect the fact that regions with cheap labor or materials are attractive export platforms.

Technology transfer through FDI. Even if they are not innovative themselves, multinational (MNC) affiliates in LDCs may transfer expertise to locally held firms through the same diffusion channels I mentioned in connection with exporters. Indeed, FDI does seem to bring relatively efficient technologies.

---

55 This result is reported in Aw and Hwang (1995); Aw and Geeta Batra (1998); Tain-Jy Chen and Tang (1987); Haddad (1993); Handoussa, Nishimizu, and Page (1986); Tybout and Westbrooke (1995); and Aw, Chen, and Roberts (1997). See, however, Rajeeva Sinha (1993).


57 Aart Kraay (1997) and Arne Bigsten et al. (1997) also find that firms become more efficient relative to others after becoming exporters in China and sub-Saharan Africa, respectively. However, because of data limitations, their econometric models are relatively restrictive. In particular, since the auto-regressive process generating average costs is constrained to be first-order, the effects of more distant cost lags may be coming through their lagged exporting dummies. (Clerides, Lach, and Tybout 1998 find that cost processes are second or third order in Morocco and Colombia.)

58 A different kind of productivity spillover from exporters occurs if their activities ease the way for other firms to break into foreign markets. Demonstration effects and the development of specialized support services like port facilities and intermediaries are possible reasons this might occur. Aitken, Gordon Hanson, and Harrison (1997) and Clerides, Lach, and Tybout (1998) report some evidence that this phenomenon is present in Mexico and Colombia, respectively.
into host countries: most studies find that foreign-owned firms are more productive than their domestically owned competitors (Haddad and Harrison 1993; Sinha 1993). But it is unclear how extensively these technologies diffuse among domestically owned firms. On the one hand, case studies suggest that substantial diffusion occurs (Blomstrom and Kokko 1997). Further, firms in sectors with relatively high MNC presence tend to be more productive in Uruguay, Mexico, Morocco, and Venezuela (Kokko, Ruben Tansini, and Mario Zejan 1997; Haddad and Harrison 1993; and Aitken and Harrison forthcoming). On the other hand, when industry effects are controlled for with dummy variables, domestically-held Venezuelan firms actually do worse as the MNC presence in their industry increases (Aitken and Harrison forthcoming). Hence cross-sectional studies may suffer from simultaneity bias because MNCs are attracted to profitable sectors, and negative spillover effects may occur in the short run because MNCs siphon off domestic demand and/or bid away high quality labor when they set up shop in the host country (Aitken and Harrison forthcoming).

Learning-by-doing and learning spillovers. As already discussed, theory tells us that protection may facilitate productivity growth by promoting domestic production in the learning-intensive sectors. Is this argument for protection empirically relevant? In developing countries, technology acquisition often amounts to adapting existing methods to local circumstances (Evenson and Westphal 1995). Hence, instead of focusing narrowly on R&D or technology purchases, the rate at which firms generate knowledge may be better proxied by the intensity with which they rely on engineers, technicians, and scientists—hereafter ETS employees. If protection encourages learning and productivity growth, one would thus expect that it helps ETS-intensive firms, and that these firms exhibit rapid productivity growth and/or generate positive spillovers.

However, it is not obvious that productivity growth and learning spillovers are greater among import-competing manufacturers than among nontradable goods producers or export-oriented producers. Arguably, the best documented case of spillovers in LDCs is the Green Revolution in Indian agriculture. Further, within each industry, the firms that export—and thus the firms that benefit from openness—tend to be more skill-intensive than others (Geeta Batra and Hong Tan 1997; Ana Revenga and Claudio Montenegro 1997; Clerides, Lach, and Tybout 1998).59

The presumption that ETS-intensive firms exhibit the most rapid efficiency growth is also tenuous. Firms with high ETS intensity do tend to get more output per unit bundle of capital and labor (Page 1980, 1984; Little, Mazumdar, and Page 1987; Cortes, Berry, and Ishaq 1987; Biggs, Shah, and Srivastava 1995). But the fact that ETS workers are more productive need not signal relatively rapid learning-by-doing, much less spillovers from one firm to another. In fact, in Colombia and Morocco, ETS-intensive firms do not exhibit higher productivity growth than others (Julie Hunt and Tybout 1997).

Finally, although common sense and case studies tell us that learning-by-doing among domestic firms is important, the available evidence suggests that it complements, rather than substitutes for, access to the international knowledge stock (Evenson and Westphal

59 This is true despite the fact that their marginal production costs tend to be lower, so it appears that highly efficient firms hire the most skilled workers and, because they are efficient, they also stand to gain the most from participation in foreign markets.
In sum, the case for fostering growth by protecting learning industries seems weak.

6. Summary

The manufacturing sectors of developing countries have traditionally been relatively protected. They have also been subject to heavy regulation, much of which is biased in favor of large enterprises. Accordingly, it is often argued that manufacturers in these countries perform poorly in several respects: (1) markets tolerate inefficient firms, so cross-firm productivity dispersion is high; (2) small groups of entrenched oligopolists exploit monopoly power in product markets; and (3) many small firms are unable or unwilling to grow, so important scale economies go unexploited.

The proliferation of very small plants and the large market shares of big plants in LDCs are sometimes interpreted to support this position. However, these distinctive features may simply trace to the general economic environment. Small geographically diffuse markets and a demand mix skewed toward simple consumer goods lead naturally to large numbers of small plants and to high concentration ratios. Indeed, although the issue remains open, the existing empirical literature does not support the notion that LDC manufacturers are relatively stagnant and inefficient. Turnover rates in plants and jobs are at least as high as those found in the OECD, and the amount of cross-plant dispersion in measured productivity rates is not generally greater. Also, although small-scale production is relatively common in LDCs, there do not appear to be major potential gains from better exploitation of scale economies.

In many countries, therefore, the main manufacturing sector problems may not be of the variety that keeps firms small, inhibits entry and exit, and/or creates market power. Rather, uncertainty about policies and demand conditions, poor rule of law, and corruption may be the priority areas for reform. These are certainly the areas that managers identify as most problematic in qualitative surveys (Brunetti, Kisunko, and Weder 1997). Also, for those countries that have not already done so, the removal of barriers to trade is likely to improve efficiency. Falling price-costs mark-ups and rising productivity have accompanied trade liberalization episodes in many LDCs.

As for future research, progress on a number of fronts would seem especially useful. First, given the fundamental importance of efficiency and productivity growth, improvements in the way that we measure these concepts should be a priority. For lack of detailed price information, most of the work thus far has equated physical output with real revenue, thereby blurring the distinction between technical efficiency and profitability. Comprehensive reckonings of the costs of productivity gains are also needed. The present value of training costs, worker recruiting and retention expenditures, technology purchases, and research programs are seldom tallied up and weighed against the present value of the productivity gains they generate. But only this kind of calculation reveals firms’ economic efficiency.

Second, given the central role that endogenous growth models assign to spillovers, any improvement in our understanding of their form and magnitude should help us to chip away at the mystery of growth. This will require better productivity measurement, as discussed above, and it will probably mean tracking individual firms over periods of time long enough to deal with impact lags. To the extent that technology diffusion takes place through labor...
turnover, data sets that merge households’ responses with those of their employers would also be useful. Progress in this arena is likely to be gradual and painful.

Third, studies that link firms’ behavior to uncertainty in the policy regime and the macro environment are scarce, given the importance that LDC entrepreneurs attach to these phenomena. In a similar vein, models that link labor market regulations and regime volatility to entry, exit, and productivity growth would address some major unanswered questions. Rigorous empirical analysis of these topics is difficult because it involves solving forward-looking optimization problems in the presence of uncertainty and sunk costs, sometimes with strategic interactions among firms and potential firms. Nonetheless, recent theoretical work has laid some of the groundwork for empirical modeling (e.g., Ericson and Pakes 1995; Avinash Dixit and Robert Pindyck 1994) and estimation techniques are improving. The next decade is likely to bring real progress.

Clearly, all of these suggested directions for research require improvements in the quality of data. Better measurement of inputs (including training, R&D, and other nontraditional factors), outputs, and prices are needed if we are to have much confidence in findings on plant-level productivity measures, or simply to document the incentive structure firms face at the ground level. More attention to data comparability across countries would also be welcome. Unfortunately, the returns to data collecting and cleaning are very small because these activities do not demonstrate cleverness to the economics profession in any obvious sense. (They are sometimes interpreted to demonstrate the opposite.) Thus database building is generally underfunded, and in cases where the investments have been made, the results have sometimes been jealously guarded rather than disseminated for widespread analysis.

REFERENCES


Pursell, Garry. 1990. “Industrial Sickness, Primary
Tybout: Manufacturing Firms in Developing Countries


