CDEP-CGEG WORKING PAPER SERIES

CDEP-CGEG WP No. 83

Firm-Level Upgrading in Developing Countries

Eric Verhoogen

March 2020
Firm-Level Upgrading in Developing Countries

Eric Verhoogen

March 2020

Abstract

In principle, firms in developing countries benefit from the fact that advanced technologies and products have already been developed in industrialized countries and can simply be adopted, a process often referred to as industrial upgrading. But for many firms, this advantage has remained elusive. What is getting in the way? This paper reviews recent firm-level empirical research on the determinants of upgrading in developing countries. The first part focuses on how to define and measure various dimensions of upgrading — increases in productivity, quality improvements, technology adoption, and expansions of product scope, among others. The second part takes stock of recent empirical evidence on the drivers of upgrading, classifying them as output-side drivers, input-side drivers, or drivers of firm capabilities. The review concludes with some thoughts about promising directions for future research in the area.

*I am grateful to David Atkin, Maria Bas, Paula Bustos, Garth Frazer, Álvaro García-Marin, Juan Carlos Hallak, Rema Hanna, Morgan Hardy, Jonas Hjort, Beata Javorcik, Seema Jayachandran, Amit Khandelwal, Gianmarco León, Florian Mayneris, David McKenzie, Pepita Miquel-Florensa, Ben Olken, Matthieu Teachout, Nico Voigtlander, and Chris Woodruff for helpful comments. I remain responsible for any errors.

‡Columbia University, eric.verhoogen@columbia.edu.
1 Introduction

At least since Gerschenkron (1962), the “advantages of backwardness” — above all, the accumulation of advanced technologies and products in industrialized countries that developing-country firms can then adopt — have been well appreciated. Since Gerschenkron’s landmark study, a number of developing countries, disproportionately in East Asia, have been able to industrialize, and to do so more quickly than earlier industrializers. But for many other countries, the purported advantages of backwardness have remained elusive. Something seems to be getting in the way of the adoption of advanced technologies and products, a process often referred to as industrial upgrading. What are these barriers? Since to identify a barrier is implicitly to identify a factor that promotes upgrading (if only by removing or mitigating the barrier), the question can be restated in a positive way: What are the drivers of industrial upgrading in developing countries?

This paper reviews recent empirical research on firms that sheds light on this question. I focus primarily on studies of larger firms (with more than a handful of employees) in manufacturing. This choice reflects a number of judgments: that such firms, although they make up a small share of the total firm population in most countries (Tybout, 2000; Hsieh and Olken, 2014), are crucial for growth; that the issues facing them are distinct from those facing very small firms, agricultural producers, and service-sector firms; and that the literatures on small firms (including entrepreneurship) and agricultural producers have been well covered by other recent reviews.\footnote{See McKenzie and Woodruff (2013), Woodruff (2018), Quinn and Woodruff (forthcoming), Foster and Rosenzweig (2010), Jack (2013), de Janvry et al. (2017), and Magruder (2018).}

To keep the review to a manageable size, I concentrate primarily (but not exclusively) on studies employing quasi-experimental and experimental methods to isolate causal relationships of interest. I also focus on studies that consider upgrading (in one of the senses discussed below) as an outcome.

The first part of the review (Section 2) discusses what is meant by the term upgrading, conceptually and empirically. The term encompasses innovation as commonly defined, but also reflects the fact that innovative activities among developing-country firms are often oriented toward catching up to the world frontier, rather than pushing it forward. I set out a simple organizing framework, which helps to clarify the four related but distinct ways in which the term upgrading has typically been used: learning, the accumulation of knowledge about products or techniques or about how to implement those techniques; quality upgrading, an increase in the weighted-average quality of goods produced by a firm; technology adoption, the adoption of a technique not previously used by a firm; and product innovation, the enlargement of the set of varieties produced by a firm. With these conceptual definitions in hand, I then review the various ways that researchers have sought to measure upgrading. As we will see, the mapping between the conceptual definitions and the empirical measures is less than perfect and existing measures have different strengths and weaknesses.

The second part of the review (Section 3) considers the drivers of upgrading. I classify them into three categories: drivers on the output side, including consumer preferences and the degree of competition in export and domestic markets (Section 3.1); drivers on the input side, including conditions in credit, labor, and intermediate-input markets (Section 3.2); and drivers of firm capabilities, including mechanisms that affect the entrepreneurial ability or knowledge possessed by firms (Section 3.3). The categorization is necessarily somewhat loose, because some mechanisms span more than one category.

A number of themes emerge from the review. First, a methodological point: there is great benefit to using directly observable information on upgrading outcomes — technology use (in-
cluding management practices), quality ratings, product scope, and productivity under controlled conditions. These measures are often available only for specific sectors, and questions naturally arise about external validity, but the approach of building up from direct observation of particular sectors seems particularly promising. Second, there is accumulating evidence that the demand side matters: selling to richer buyers, or supplying inputs in value chains destined for richer buyers, seems to be robustly associated with upgrading. Third, evidence has also accumulated of causal links between the cost, quality, and variety of inputs and upgrading outcomes. Increased access to imported inputs, for instance, appears to stimulate upgrading. Fourth, it is clear that developing-country firms are often constrained by a lack of know-how. Several types of informational interventions have been successful in improving firm performance. At the same time, organizational dynamics are complex and learning is costly, and a lack of upgrading should not simply be attributed to a failure of individuals to optimize. A number of other insights will be highlighted as we proceed.

This review is related to a number of existing reviews, beyond those cited above. In its focus on firms in developing countries, it is similar in spirit to an older review by Tybout (2000), but with different topical emphases. Several reviews from the perspective of international trade have covered work in developing countries, including Tybout (2003), De Loecker and Goldberg (2014), Goldberg and Pavcnik (2016), Shu and Steinwender (2019), and Atkin and Khandelwal (2019); the current review is broader in considering drivers of upgrading unrelated to trade, but also narrower in focusing on firm-level empirical work on upgrading outcomes using quasi-experimental and experimental strategies. Also related are the handbook chapter of Harrison and Rodriguez-Clare (2010) on the theory and practice of industrial policy in developing countries, and recent policy-oriented overviews by Crespi et al. (2014), Cirera and Maloney (2017), and Cusolito and Maloney (2018). The current review is focused on evaluating what we know about how firms behave, which is relevant to policy design, but not specifically on the practical issues of what works or does not work in industrial policy.

2 What Do We Mean by Upgrading?

The word upgrading is used in a variety of ways. This section aims to clarify, conceptually and empirically, how the term has been used and to highlight the strengths and weaknesses of existing empirical measures.

2.1 A Simple Framework

To organize the discussion, some notation and a simple, general framework will be useful. We can think of a firm, indexed by $i$, as a collection of production lines each producing a single product, indexed by $j$, using one production technique, $k$, at time $t$, characterized by a product-technique-specific production function:

$$Y_{ijkt} = F_{ijk}(\tilde{M}_{ijkt}, \lambda_{ijkt})$$

where $Y_{ijkt}$ is physical output, $\tilde{M}_{ijkt}$ is a vector of physical inputs (which may include outputs from other production lines in the firm) and $\lambda_{ijkt}$ is what Sutton (2007, 2012) and others call the capability of firm $i$ in product-technique $jk$, which is assumed to raise output conditional on

\footnote{See also Lane (forthcoming).}
inputs (i.e. \( \frac{\partial Y_{ijkt}}{\partial \lambda_{ijkt}} > 0 \)). The set of capabilities can also be understood to incorporate what Dessein and Prat (2019) term “organizational capital,” a firm-specific asset that must be produced within the firm and changes slowly over time. A technique can be thought of as a set of instructions for combining particular machines and practices with particular inputs. Let \( \Lambda_{it} \equiv \{\lambda_{ijkt}\} \) be the set of capabilities of a firm and let \( J_{it} \) and \( K_{ijt} \) be the sets of products and corresponding techniques for which the firm knows \( F_{ijk}(\cdot) \).

To keep language simple, I will refer to \( \Lambda_{it}, J_{it}, \) and \( K_{ijt} \) together as “know-how.”

Suppose that \( P_{ijt} \) is the firm’s output price for product \( j \), and that the output demand curves facing the firm are given by \( P_{ijt} = D(Y_{ijt}, \tilde{Y}_{i,-jt}, Z^D_t) \), where \( Z^D_t \) reflects external factors in the output market. Similarly, suppose that the vector \( \tilde{W}_{ijkt} \) holds prices for the inputs used in product-technique \( jk \), and that the input supply curves facing the firm are given by \( \tilde{W}_{ijkt} = S(M_{ijkt}, \tilde{M}_{i,-jkt}, Z^S_m) \), where \( Z^S_m \) reflects external factors in input markets.\(^4\) It is also assumed that the firm faces fixed costs of production, which may be at the level of a product-technique, \( f_{ijkt} \), a product, \( f_{ijt} \), or the firm, \( f_{it} \), and which may vary across firms (and depend, for instance, on a firm’s capabilities) or across destination markets. It is also assumed that the firm can affect its future capabilities or expand the sets of products and techniques that it knows about by making investments \( I_{it}^\Lambda, I_{it}^J, \) and \( I_{it}^K \), respectively. A firm’s future know-how may also be affected by the set of products it chooses to produce, or the techniques it uses to produce them.

The firm’s present discounted profit can then be written as:

\[
\Pi_{it} = \sum_{t=0}^{\infty} \delta_t \left\{ \sum_{j \in J^*_{it}} P_{ijt} F_{ijk^*}(\tilde{M}_{ijk^*t}, \lambda_{ijk^*t}) - \tilde{W}^{ijkt}_{ijk^*t} \tilde{M}_{ijk^*t} - f_{ik^*t} - f_{ijt} \right\} - f_{it} - I_{it}^\Lambda - I_{it}^J - I_{it}^K
\]

where \( \delta_t \) is a discount factor, \( J^*_{it} \) is the set of products the firm chooses to produce, and \( k^*_{ijt} \) (indicated by the \( k^* \) subscript) is the optimal technique chosen for each product, \( j \in J^*_{it} \). The firm’s decision problem in any period is to choose \( J^*_{it}, k^*_{ijt} \) for each \( j \in J^*_{it} \), the amount of each input used for the chosen product-technique, \( \tilde{M}_{ijk^*t} \), and investments in future know-how, \( I_{it}^\Lambda, I_{it}^J, \) and \( I_{it}^K \), in order to maximize the firm’s present discounted profit, \( \Pi_{it} \).\(^5\)

In its current form, the framework is too general to be able to generate falsifiable predictions about firm behavior, but it is helpful to define terms and to organize our thinking. The most common definitions of upgrading in the literature can be classified conceptually under four headings, which I will refer to as learning, technology adoption, quality upgrading, and product innovation. These dimensions are related and often occur together but are conceptually distinct.

We can think of learning as an accumulation of know-how: an increase in capabilities, \( \lambda_{ijkt} \), for some subset of product-techniques, an expansion of the set of products the firm knows about, \( J_{it} \), or an expansion of the set of techniques the firm knows about for a given product, \( K_{ijt} \).

\(^{3}\)To keep things simple, I assume that a firm either knows \( F_{ijk}(\cdot) \) or not, i.e. that there is no partial knowledge of techniques. In reality, a firm might have uncertainty about \( F_{ijk}(\cdot) \) and reductions in such uncertainty are an important component of learning. See e.g. Foster and Rosenzweig (2010).

\(^{4}\)Again, to keep things simple, I assume that the firm knows the demand and supply functions, \( D(\cdot) \) and \( S(\cdot) \), but in reality a firm may have imperfect knowledge and may invest in learning about these relationships.

\(^{5}\)There may be adjustment costs involved in changing products or techniques, which can be captured in this framework by the (potentially time-varying) fixed costs \( f_{ijkt}, f_{ijt}, \) and \( f_{it} \). Note that the firm does not necessarily optimize on each production-line independently; for various reasons, including capacity constraints because of fixed factors such as entrepreneurial attention, choices on one line are likely to affect choices on others.
in the framework is a distinction between skills that can be purchased on the labor market (and hence show up in $M_{ijkt}$) and capabilities and knowledge that must be acquired through other means, which may include conscious investments ($I^A_{it}$, $I^I_{it}$, and $I^K_{it}$) or incidental learning from one’s own experience or the experiences of others. Learning would certainly include expansions in $J_{it}$ or $K_{ijt}$ that also expand the sets of products and techniques available to the world, denoted as $J_t$ and $K_{jt}$, but as mentioned above this sort of new-to-the-world innovation (as opposed to new-to-the-firm innovation) is rare in developing countries.

*Technology adoption* can be thought of simply as the employment of a technique not previously in use by the firm. Here I will use a broad definition of techniques that includes management practices; these are considered to be chosen by firms, given their capabilities. In this framework, production processes are components of techniques, and process innovation can be considered a form of technology adoption. It is tempting to limit the definition of technology adoption to adoption of technologies that are in some sense better than the technologies a firm is currently using. The difficulty here is that technologies are rarely “better” in a global sense — that is, better for all possible levels of know-how and output-demand and input-supply functions. Empirically, it is almost never possible to establish that technologies are globally superior in this way. I will therefore maintain the more agnostic definition of technology adoption as adoption of any technique not previously used by the firm.

Before defining quality upgrading, we need to be clear about what we mean by product quality. It is useful to think about demand functions such as $D(\cdot)$ above as summarizing demands from individual consumers with heterogeneous preference draws (see e.g. Anderson et al. (1992)). A product can be considered to be of higher quality than another in the same product category if it has a higher market share when priced at the same level, although some individuals will be idiosyncratically attached to each product. Quality can be thought of as a one-dimensional index of product attributes that predicts market share conditional on price. In this framework, we can think of output varieties of different qualities (within a product category) as simply being different products, with different labels $j$. Similarly, inputs of different qualities can be thought of as forming part of different techniques. Quality upgrading can then be defined as an increase in the average quality of goods produced, as reflected in an output-weighted average of the goods in $J^*_{it}$. A firm can upgrade in this way without producing goods new to the firm, by shifting output toward higher-quality products already being produced.

*Product innovation* can be thought of as the production of a good not previously produced by a firm. Product innovation does not necessarily involve learning as defined above, since a firm may start producing a product that is already in the set of products it knows about, $J_{it}$, that it happened not to produce before. Product innovation is also distinct from quality upgrading, since the new products may or may not be of higher quality than the products already being produced. Product innovation may entail the switch of a firm to a new sector (given the common practice of assigning firms to the sectors in which they have a plurality of their sales) but most cases do not involve such sectoral shifts.

---

6 In treating management practices as technologies, I am following, among others, Van Reenen (2011), who argues that the choice of management practices should be analyzed as one would analyze any other technology choice, and Bloom et al. (2011), who write, "Modern management is a technology that diffuses slowly between firms." See also Bloom et al. (2017).

7 That is, production processes that use the same sets of machines or practices but different qualities of inputs would be considered different techniques.

8 A reasonable alternative definition of quality upgrading would be an increase in the highest quality product produced by a firm. As a practical matter, average quality and maximum quality are likely to be highly correlated.
This framework motivates the categorization of drivers of upgrading in Section 3 below. One set of drivers has to do with conditions in output markets, here summarized by the demand curves, \( D(Y_{ijt}, \bar{Y}_{i,-jt}, Z^y_t) \). Another set of drivers has to do with conditions in input markets, here summarized by the input-supply curves, \( S(\bar{M}_{ijkt}, \bar{M}_{i,-jkt}, Z_{mt}) \). A third set has to do with the know-how of firms, here summarized by \( \Lambda_{it}, J_d \) and \( K_{ijt} \). The demarcation between categories is not sharp. For instance, firms’ capabilities may shape their decisions about which output markets to enter and which sets of consumers to face. Similarly, output or input market conditions, by influencing which products firms produce and which techniques they use, may affect how quickly firms learn. But the categorization seems to be a reasonable way to organize existing studies.

In addition to helping to define terms, this framework highlights three key conceptual points. First, the conditions facing entrepreneurs in developing countries typically differ in a number of ways from those facing firms in developed countries. Developing-country firms often face different (typically poorer) consumers and different prices in input markets, and they have different levels of know-how. These factors influence firms’ choices of which products to produce and which techniques to use.

Second, the four dimensions of upgrading, as we have defined them, are not necessarily optimal for firms or beneficial for aggregate economic performance. More know-how is a good thing for firms, but if acquiring know-how is costly, a firm must weigh the required investment against the future benefits of learning. Whether producing new and/or higher-quality products, or using new techniques, is optimal will depend on conditions in output and input markets and a firm’s level of know-how. When seeking to interpret upgrading behavior, or lack thereof, researchers need to keep in mind the heterogeneous constraints and opportunities faced by firms.\(^9\)

Third, understood through the lens of this framework, the popular conception of “management” reflects three related but conceptually distinct elements: entrepreneurial ability, which we can think of as a component of capabilities, \( \lambda_{ijkt} \), that is common across products and techniques and embodied in an entrepreneur; the skill of employed managers, which can be thought of as a component of the input vectors, \( M_{ijkt} \); and the management practices chosen by the firm, which are components of the selected techniques, \( k^*_{ijt} \). In this view, it is not sufficient to attribute poor firm performance to “bad management”; one needs specify how each of these three elements play a role in the poor outcomes.\(^{10}\) We will return to these issues in Section 3.3 below.

2.2 Measurement Issues

We now turn to the question of how to measure upgrading. There is an important tension here. On one hand, the most common empirical measure, total factor productivity (TFP) in various forms, is conceptually attractive in that it is aimed directly at measuring firm capabilities, and improvements in capabilities in theory bear an unambiguously positive relationship to firm performance. But TFP measures also suffer from a number of well-known potential biases.\(^{11}\) On the other hand, direct indicators of product quality, product innovation, and technology adoption are increasingly available, and are arguably more credible measures of the dimensions of upgrading.

---

\(^9\) As Foster and Rosenzweig (2010) write in an agricultural context, “it cannot be inferred from the observation that farmers using high levels of fertilizer earn substantially higher profits than farmers who use little fertilizer that more farmers should use more fertilizer” (p. 399).

\(^{10}\) There may of course be interactions between these elements: for instance, low-ability entrepreneurs may choose low-skill managers, who in turn choose sub-optimal management practices.

\(^{11}\) Many of the issues raised below are discussed in more detail in previous reviews by Bartelsman and Doms (2000), Katayama et al. (2009), Ackerberg et al. (2007), and De Loecker and Goldberg (2014).
they seek to capture. But it is not always obvious what constitutes an “improvement” on these dimensions, and the indicators are typically only available in particular sectors, raising questions about external validity. This subsection considers the strengths and weaknesses of the different measures that have been employed in the literature.

2.2.1 Measures of Productivity

The standard approach to TFP estimation begins by positing the existence of a firm-level production function, most commonly Cobb-Douglas, for instance:

\[ y_i = \beta_k k_i + \beta\ell \ell_i + \beta_m m_i + \{\omega_i + \varepsilon_i\} \]  

(3)

where \(y_i\) is log output, typically sales deflated by a sector-level output price deflator, \(k_i\) is log capital, \(\ell_i\) is log labor, (employment or hours), \(m_i\) is log materials, typically expenditures deflated by a sector-level input price deflator, \(\omega_i\) is “ex ante” productivity, which the firm knows before choosing the variable inputs \(\ell\) and \(m\), and \(\varepsilon_i\) is an “ex post” shock, realized after the firm had made its input decisions.\(^{12}\) The coefficients \(\beta_k\), \(\beta\ell\), and \(\beta_m\) are then estimated by one of several methods (discussed briefly below), and TFP is estimated as: \(\hat{TFP}_i \equiv y_i - \hat{\beta}_k k_i - \hat{\beta}\ell \ell_i - \hat{\beta}_m m_i\).

One under-appreciated issue with this approach is that if the firm is actually a collection of production lines, as in the framework above, then it is not obvious that there exists an aggregate production function that fully summarizes the relationship between inputs and outputs at the firm level. Under certain conditions, production-line-level production functions such as the \(F_{ijk}(\cdot)\) in equation (1) aggregate into a firm-level function such as equation (3).\(^{13}\) This finding is analogous to earlier results on the aggregation of firm-level production functions to a macro-level production function, going back to Houthakker (1955). But the assumptions required in the earlier literature have been criticized as special and unlikely to hold in practice (Felipe and Fisher, 2003), and a similar point could be made about the aggregation from the firm-product-technique level to the firm level. A main defense of standard aggregate production functions has been that they seem to work pretty well, in that they provide a reasonable fit between aggregate inputs and aggregate output and the estimated factor elasticities of output are consistent with observed factor shares (Fisher, 1971; Fisher et al., 1977), and a similar defense could be made for firm-level production functions such as equation (3). But given the shaky microfoundations, caution is warranted in interpreting them. The caveat of Mairesse and Griliches (1988) still seems apt: “[T]he simple production function model ... is at best just an approximation to a much more complex and changing reality at the firm, product, and factory floor level” (p. 28).

Much of the recent literature on production-function estimation has been concerned with a different problem, the “transmission bias” recognized by Marschak and Andrews (1944): in the context of equation (3), if a firm observes that it has a high ex ante productivity, then it may choose to use more labor and materials, generating a correlation between \(\omega_i\) and \(\ell_i\) and \(m_i\) and biasing OLS estimates. The most common way to address this issue is to construct an observable proxy

\(^{12}\)This is a “gross output” production function; an alternative is to estimate a “value-added” production function; for advantages and disadvantages, see Ackerberg et al. (2015) and Gandhi et al. (forthcoming, 2017).

\(^{13}\)For instance, Jones (2005) considers an environment in which a firm produces a single product and chooses over Leontief techniques, where the Leontief coefficients are drawn from independent Pareto distributions. As the set of techniques over which the firm chooses becomes large, the maximum output for a given set of factor choices can be expressed as a Cobb-Douglas function similar to equation (3). Subsequent research has derived similar results in this spirit, with specific assumptions on functional forms and distributions of technique draws (Growiec, 2008a,b; Boehm and Oberfeld, 2018).
for the ex ante productivity term, using either investment (Olley and Pakes, 1996) or materials (Levinsohn and Petrin, 2003; Ackerberg et al., 2015). These approaches have recently been criticized by Gandhi et al. (forthcoming), who argue that the Olley-Pakes and Levinsohn-Petrin estimators are not non-parametrically identified; they propose using the first-order condition for the choice of materials as an additional source of identification. It is also important to note that the monotonicity assumption required for standard proxy-variable methods is strong; in the Olley-Pakes version, for instance, heterogeneity across firms in the extent to which they are credit constrained or face adjustment costs of capital would violate the required assumption (Griliches and Mairesse, 1998; Ackerberg et al., 2015).

A separate issue arises because it is rare to observe physical quantities of outputs or inputs. It is common is to use sector-level output and input price deflators to deflate firm-level revenues (or value-added) and input expenditures. But as De Loecker and Goldberg (2014) point out, this can give rise to potentially severe biases, if idiosyncratic factors that affect output or input prices are correlated with a firm’s input choices, as in general one would expect them to be. Datasets with physical quantities at the firm-product level are increasingly available, and in sectors with homogeneous products the quantity information can help to address these biases. In US data, Foster et al. (2008) focus on 11 arguably homogeneous products and estimate a function with physical output on the left-hand side, to yield what they call TFPQ (Q for quantity). They contrast it with a measure of TFP estimated with revenues on the left-hand side, TFPR (R for revenues). Although the US data do not contain physical quantities of inputs, such information is available in a few other countries (Chile, Colombia, Ecuador, Peru, Portugal, and Spain, among others), and one could in principle include physical inputs on the right-hand side to solve the input-price bias problem.

But it is important to be aware that quantity-based TFP measures are likely to be a misleading indicator of firm capability in the presence of quality differences in either inputs or outputs (Katayama et al., 2009; Grieco and McDevitt, 2016). Intuitively, a firm may take advantage of increased capability to raise quality rather than simply to increase physical output, leading quantity-based TFP to understate the true capability change. Differences in input quality may generate an offsetting bias, if firms are able to create more units of output out of higher-quality inputs. In the Melitz (2003)-type theoretical model of Kugler and Verhoogen (2012), these effects can arise under certain parameter values. This is not just a theoretical curiosity. In an

14 Intuitively, in the Olley and Pakes (1996) case, in the context of a value-added production function, if investment is a function of productivity and existing capital stock, \( i_t = \lambda(\omega_t, k_t) \), and \( \omega_t \) is a scalar and strictly monotonically related to \( i_t \) then this function can be inverted, and the productivity term can be expressed as a function of investment and capital: \( \omega_t = h(i_t, k_t) \). A flexible polynomial in \( i_t \) and \( k_t \) can then serve as a proxy for \( \omega_t \) in an equation similar to equation (3). Levinsohn and Petrin (2003) propose a similar approach for materials. Ackerberg et al. (2015) also invert a materials-demand equation, but (in contrast to Levinsohn and Petrin (2003)) one that conditions on labor inputs.

15 Gandhi et al. (2017) note that their criticism in Gandhi et al. (forthcoming) does not apply in a setting where a linear function of materials is a perfect complement to other inputs in producing output; this setting yields the value-added specification employed by Ackerberg et al. (2015).

16 For instance, OLS estimates will be biased if a firm faces idiosyncratically high input prices and spends less on inputs as a result (De Loecker and Goldberg (2014) call this “input-price bias”) or faces idiosyncratically high output prices and spends more on inputs as a result (“output-price bias”).

17 In Kugler and Verhoogen (2012), higher productivity leads to lower input requirements conditional on product quality but also leads firms to produce higher-quality goods, which carry a higher price. Whether physical units of output increase or decrease with firm capability depends on the elasticity of demand faced by the firm, the extent to which capability reduces unit costs conditional on quality, and the scope for quality differentiation in the industry. (In that model, physical output as a function of capability can be readily calculated by dividing revenues by output
experiment discussed at greater length below, Atkin et al. (2017a, 2019) randomly allocated export contacts to Egyptian rug producers. They find that the producers increased exports, quality (which they measure directly), and profits, as might be expected, but decreased square meters of rug woven per hour and TFPQ. In laboratory conditions, sewing identical rugs, the treated weavers were no slower than the non-treated weavers and they sewed higher-quality rugs. In this setting, it seems clear that TFPQ is misleading as a measure of firm performance. Although it may only be in extreme cases that measured TFPQ is negatively affected by increases in firm capability, we would expect quality changes to drive a wedge between TFPQ and capability — the theoretical concept one would like to measure — in a wide variety of circumstances. Quality bias of this sort is likely to be particularly salient in developing countries as firms enter world markets, because of the large differences in incomes between domestic and rich-country consumers.

A natural response to the issues of quality bias is to revert to using revenues on the left-hand side and expenditures on the right-hand side. Using price times quantity, rather than just quantity, should take into account quality differences, since they are presumably reflected in prices. But prices also reflect things other than quality, in particular markups. In imperfectly competitive industries, TFPR is a measure both of technical efficiency — the ability to transform physical inputs into physical outputs — and of the ability to sell at a price above marginal cost (De Loecker and Goldberg, 2014). It may well be the best measure of firm performance available for quality-differentiated industries, but one should not interpret it solely as a measure of technical efficiency. One way to address this issue is to estimate markups directly to separate them from marginal costs (which reflect technical efficiency); we return to this issue below.

When estimating productivity with the new data on physical quantities, one must also decide whether and how to aggregate across products in multi-product firms. Even datasets with product-level information typically do not report which inputs are used to produce which outputs. One approach is to focus on single-product firms and possibly to do a selection correction for the fact that they are not representative (Foster et al., 2008; De Loecker et al., 2016; Balat et al., 2018). Another is to impose theoretical structure on the demand side and to use the model to infer how firms would allocate inputs to outputs if they were behaving optimally (Orr, 2018; Valmari, 2016). The literature has not yet converged on a consensus approach to this issue.

In sum, although TFP measures have the attractive property that they aim directly at estimating firm capabilities, existing estimation methods suffer from a number of well-known difficulties and may reflect a number of other factors besides capabilities — notably markups in the case of TFPR and endogenous quality choices in the case of TFPQ. We will see below that results for TFP outcomes are often mixed. This may in part be due to a confounding of effects on firm capabilities with effects on markups or quality choices.

### 2.2.2 Measures of Quality

Direct measures of quality are not available in standard firm-level datasets and are typically quite difficult to come by. But a few studies have had access to direct information on firm-level quality

---

18 In another illustration, De Loecker et al. (2016) pursue a more model-based approach to estimating production function parameters, allowing for quality differences on both the input and output sides. They find plausible estimates when they control for quality differences, but nonsensical estimates when they do not (Table V). See the further discussion in Sections 3.1.1.3 and 3.2.1 below.

19 Another example, of kidney dialysis centers in the US, is provided by Grieco and McDevitt (2016).

20 The two exceptions I am aware of for a large number of firms are the dataset on the Bangladeshi garment sector used by Cajal Grossi et al. (2019) and the dataset on Chinese steel firms used by Brandt et al. (2018).
choices. Several recent papers have used quality ratings (or prizes at tasting competitions) for wines, in France (Crozet et al., 2012), Chile (Macchiavello, 2010), and Argentina (Chen and Juvénal, 2016, 2018, 2019). Studies have taken advantage of direct information on quality of Egyptian rugs (Atkin et al., 2017a), sweetness of watermelons (Bai, 2018), contamination of dairy products (Bai et al., 2017), automobile defects (Bai et al., 2019), the protein content of fishmeal (Hansman et al., forthcoming), and coffee bean characteristics such as size and defect rates (Macchiavello and Miquel-Florensa, 2018, 2019). Verhoogen (2008) proxies for quality using ISO 9000 certification, an international production standard. Accessing more direct measures of quality to examine firm-level quality choices is a promising direction for research.

An alternative approach is to construct measures of quality from information on prices and quantities, which requires theoretical structure. Khandelwal et al. (2013) show how this can be done in trade-transactions data on Chinese textile and clothing firms. In a Melitz (2003)-type model where a representative consumer has CES preferences and values product quality, the product-level demand functions facing a firm can be written as: $\ln Y_{ijt} = -\sigma \ln P_{ijt} + \alpha_j + \alpha_t + \epsilon_{ijt}$, where $Y_{ijt}$ is product quantity, $P_{ijt}$ is price, $\sigma$ is the elasticity of substitution between products, $\alpha_j$ and $\alpha_t$ are product and year fixed effects, respectively, and $\epsilon_{ijt}$ equals quality times $\sigma - 1$. The authors set $\sigma = 4$, the median elasticity of substitution for clothing and textile products from Broda et al. (2006), and rewrite the expression as $\ln Y_{ijt} + \sigma \ln P_{ijt} = \alpha_j + \alpha_t + \epsilon_{ijt}$. They run this regression, recover the residual $\hat{\epsilon}_{ijt}$, and interpret $\frac{\hat{\epsilon}_{ijt}}{\sigma - 1}$ as a measure of quality at the firm-product level. The intuition is the same as discussed in Section 2.1 above: conditional on price, higher quality products have higher market share and hence a higher $\hat{\epsilon}_{ijt}$. This method is akin to methods to recover quality at a more aggregate level by Hummels and Klenow (2005), Khandelwal (2010), Hallak and Schott (2011), and Feenstra and Romalis (2014), among others. Variations have been used by Bas and Strauss-Kahn (2015), Fan et al. (2015, 2018), Stiebale and Vencappa (2018), and Bas and Paunov (2019).

While the Khandelwal et al. (2013) method has proven useful, it requires several non-innocuous assumptions, both in the specification of demand and in the estimation of $\sigma$ carried out by Broda et al. (2006). An alternative approach uses reduced-form relationships between prices and other observables to argue indirectly that quality differences appear to be playing an important role, without imposing the functional form assumptions required to construct explicit measures of quality. Kugler and Verhoogen (2012) take advantage of rich data from the Colombian manufacturing census on output and input prices to document several facts. First, on average within narrow product categories, larger plants charge higher prices for their outputs. Second, larger plants also pay more for their material inputs — a fact that generalizes the well-known finding in labor markets that larger firms tend to pay higher wages (Brown and Medoff, 1989). Third, the output price-plant size and input price-plant size correlations are more positive in sectors with greater scope for quality differentiation, where, following Sutton (1998), the scope for quality differentiation is proxied by R&D and advertising expenditures. The empirical patterns are difficult to reconcile with models that do not accord an important role to quality differences and suggest that producing high-quality outputs requires high-quality inputs, a hypothesis that has been corroborated by other studies discussed below.

An important caveat is that one should be cautious about interpreting high prices alone as

---


22 Khandelwal et al. (2013) observe prices and quantities separately by export destination in Chinese customs data, and include a destination-year fixed effect.
indicators of quality, even if they are correlated with high input prices. Firms may face positive input cost shocks, and they may pass those on to consumers in the form of high prices. But in the absence of quality differences, we would expect such high-cost firms to have smaller market shares than low-cost firms. This underlines the need to examine sales (or other indicators of firm size) in addition to prices before drawing strong conclusions about quality.

2.2.3 Measures of Technology Use

Direct information on technologies used by manufacturing firms is also often difficult to obtain. Standard firm-level datasets do not contain it, and firms are often reluctant to speak about specific technologies, for fear of revealing proprietary information to competitors. The technology-adoption literature has tended to focus on agriculture, where information on technology use is more readily available (Foster and Rosenzweig, 2010). In developed countries, there have been a number of studies of technology adoption across reasonably large sets of manufacturing firms, for instance the “insider econometrics” studies reviewed by Ichinoewski and Shaw (2013), and studies of adoption of energy-efficient technologies reviewed by Allcott and Greenstone (2012). In developing countries, studies employing direct measures of technology use by manufacturing firms have been scarcer, but include the recent papers on Pakistani soccer-ball producers by Atkin et al. (2017b) and on Ghanaian garment producers by Hardy and McCasland (2016), which we discuss in a later section. The World Bank is currently engaged in a series of surveys of technology use in developing countries, which are likely to stimulate increasing work in the area. One challenge in this line of research is that machines and other physical technologies are often specific to particular sectors and can only be captured by detailed, tailored surveys. Also, as noted above, it is often unclear the extent to which one technology can be considered “better” than another. But measures of technology use, when available, have the great advantage that they are informative even in the absence of strong functional-form assumptions.

As discussed above, we can think of management practices as a form of technology. The measurement of management practices has been advancing rapidly, following the influential work of Bloom and Van Reenen (2007, 2010). The World Management Survey (WMS) was first implemented in the US and Europe but has now been extended to 35 countries, including low-income countries such as Ethiopia and Mozambique (Bloom et al., 2014). Using open-ended questions on monitoring, production targets, and incentives, posed by skilled interviewers, the survey has constructed management scores that have proven to be robustly correlated with a variety of independent measures of firm performance. Information on management practices has also been collected using “closed-ended” (i.e. multiple-choice) questions in the Management and Organizational Practices Survey conducted by the US Census and in similar surveys in Mexico, Pakistan, and other countries (Bloom et al., 2016b, 2019).\footnote{Relatedly, McKenzie and Woodruff (2017) review findings from seven countries using a battery of questions designed for smaller developing-country firms.} An important advantage of focusing on management practices as a form of technology use is that similar practices are applicable across a wide range of contexts. It has been possible to construct consistently measured management scores across a range of countries and sectors, and this in part explains the substantial impact of this research agenda on several fields.

There is a debate in this literature about whether particular practices can be considered better than others in some absolute or context-independent sense. On one hand, there is a long tradition in management research, often referred to as the “horizontal” (or “design” or “contingency”) view, that sees the best management practices as contingent on many features of a firm’s environment...
On the other hand, the key proponents of this literature argue for a “vertical” view that some practices are better than others across settings (see e.g. Van Reenen (2011) and Bloom et al. (2014)). This is ultimately an empirical question, one that in my view is not yet resolved. As with other technologies, one should not infer from the mere fact that more-successful firms use a particular practice that all firms should adopt it. Firms may lack the know-how to implement the practice effectively, or may face different output market or input market conditions than those who use the practice successfully. It seems likely that some firms are making mistakes by not adopting some higher-scoring practices (e.g. tracking inventories). But for other practices (e.g. performance pay) the situation is less clear-cut. It seems important to consider carefully firms’ capabilities and the settings in which they operate before concluding that a particular practice is better than another.

2.2.4 Measures of Product Innovation

The most common measures of innovation-related activities in developed countries are patents and R&D expenditures. But as discussed above, most innovation-related activities in developing countries are directed towards catching up to the world frontier, not extending it, and such efforts are typically not reflected in patents or R&D (although there have been a few studies, some of which are reviewed below). An arguably more informative approach for developing countries is to focus on the range of products produced by a given firm. This is increasingly feasible as firm-product-level datasets become more widely available. As data at the firm-product level become increasingly available, it is becoming possible to observe product innovation directly, as additions to the set of products produced by a firm (see e.g. Goldberg et al. (2010), Bas and Paunov (2019).) Access to barcode-level product data, linkable to firms, is expanding rapidly in developed countries (e.g. Faber and Fally (2017)) and developing countries (e.g. Atkin et al. (2018)), and incorporating this rich new information would be a promising direction for research.

2.2.5 Discussion

There are costs and benefits to each of the measures of upgrading we have considered. TFP measures aim most directly at estimating a firm’s capabilities, $A_{it}$, which in theory are unambiguously related to technical efficiency and firm performance. But the difficulties in TFP estimation are many, and, perhaps as a consequence, results with TFP as an outcome have been mixed. The other indicators we have considered are often available only in specific settings and need to be interpreted with caution, since it is not obvious that increases in them are optimal for firms or beneficial for growth, but they typically require fewer auxiliary assumptions. It seems clear that the literature should continue to consider various measures of upgrading, and that we should have the most confidence in patterns that show up consistently across measures. But beyond that, my sense is that the most compelling recent studies are those that have focused on directly observable measures, and that expanding the settings in which such information is available is a promising avenue for research.

---

24 For example, Bloom et al. (2014, p. 852) write, “The focus of the WMS questions is on practices that are likely to be associated with delivering existing goods or services more efficiently. We think there is some consensus over better or worse practices in this regard.”
3 Drivers of Upgrading

We now turn to our central question: what are the drivers of upgrading? I categorize drivers into three groups, which can be understood with reference to the general framework above: (1) output-side drivers: factors that affect product demand curves (the $D(\cdot)$ functions); (2) input-side drivers: factors that affect input-supply curves (the $S(\cdot)$ functions); and (3) drivers of capabilities: factors that affect the the “know-how” of firms (the $A_{ijkt}$, $J_{it}$, and $K_{ijkt}$). This categorization is necessarily somewhat loose — some drivers fit in more than one category, and some not quite in any — but the grouping is helpful to organize the review.

3.1 Output-Side Drivers

We begin with the literature on the effects of exporting on upgrading outcomes, because the literature is perhaps the most fully developed, and then turn to other output-side drivers, including the effects of local demand from multinational enterprises, competition in output markets, and other factors.

3.1.1 Exports

Early studies on exporting and productivity — Bernard and Jensen (1995, 1999) using US data and Clerides et al. (1998) using Mexican, Colombian, and Moroccan data — find little evidence that firms increase productivity when they start exporting. Instead, the superior performance of exporters in cross-section is explained by the selection of already-higher-performing firms into exporting. The influential Melitz (2003) model was written with these results in mind and is consistent with them: under monopolistic competition and trade between symmetric countries, firms with a sufficiently high initial productivity draw enter the export market, but increases in exporting have no within-rm effects on productivity, output quality, or wages. More recent evidence, however, has found robust effects of exporting on a number of upgrading outcomes.

3.1.1.1 Exports and Quality

A first-order feature of the world economy, from the perspective of manufacturing firms in developing countries, is that consumers in international markets are on average richer and more willing to pay for product quality than domestic consumers.\textsuperscript{25} A natural corollary is that a given firm in a developing country will produce higher-quality goods for export to rich countries than for sale in its own domestic market, to appeal to richer consumers. Verhoogen (2008) develops this idea in a Melitz (2003)-type heterogeneous-firm framework.\textsuperscript{26} In

\textsuperscript{25}In trade, the idea that consumers in richer countries are more willing to pay for quality is commonly attributed to Linder (1961). In the consumption literature, the idea is regarded as so well established as to be unremarkable; see e.g. Deaton and Muellbauer (1980).

\textsuperscript{26}Several earlier empirical papers explore the role of quality in trade at a more aggregate level. In addition to Hummels and Klenow (2005), cited above, Schott (2004) shows that the US imports higher-priced products within narrow trade categories from richer countries, suggesting quality differences. In a cross-country setting, Hallak (2006) shows that richer countries tend to demand relatively more from exporters with higher prices (and presumably higher quality). Notable early theoretical papers on quality in trade include Gabszewicz et al. (1982) and Flam and Helpman (1987). It appears that Verhoogen (2008) was the first to use a heterogeneous-firms model to formalize the idea that a given firm will sell a higher-quality variety in a richer market and to explore its implications in firm- (or plant-) level data. The related but distinct idea that firms’ quality choices respond to per-unit trade costs (as in the famous example of Washington apples from Alchian and Allen (1964)) has been developed by Rodriguez (1979), Feenstra (1988), Hummels and Skiba (2004), Feenstra and Romalis (2014) and others.
addition to non-homotheticity of consumer demand, the key theoretical supposition is that firm capability and input quality are complementary in producing output quality. It follows that more-capable firms use higher-quality inputs to produce higher-quality outputs in equilibrium. As in Melitz (2003), only more-capable firms enter the export market. An exogenous increase in the incentive to export leads plants that are already exporting to shift production toward higher-quality varieties and induces some firms that are not exporting to enter the export market. Average product quality and hence average input quality and average wages increase in more-capable firms relative to less-capable firms.\textsuperscript{27} Empirically, the paper tests this prediction at the plant level using initial plant size as a proxy for capability (since more-capable plants grow to be larger) and examining the differential response of Mexican plants to the late-1994 peso devaluation. Initially larger plants increased exports, were more likely to acquire ISO 9000 certification (an international production standard interpreted as a proxy for product quality), and increased wages relative to initially smaller plants within the same industry. The differential response was not evident in periods without devaluations. The differential quality upgrading generates a link between trade and wage inequality, since the initially larger plants already paid higher wages and further increased wages relative to initially smaller plants within industries.\textsuperscript{28}

This basic story has held up reasonably well and has been extended by subsequent research. One source of evidence is price correlations in more disaggregated data. Using trade-transactions data from customs agencies, several papers have documented that firms charge higher prices in richer destinations within narrow product categories. Bastos and Silva (2010) first documented this pattern in Portuguese data, and it has been shown to be robust in Chinese (Manova and Zhang, 2012), French (Martin, 2012), and Hungarian (Görg et al., 2017) data. As mentioned above, Kugler and Verhoogen (2012) document positive correlations between output prices, input prices, and plant size that suggest producing high-quality outputs requires high-quality inputs, consistent with a general-equilibrium model similar to Verhoogen (2008).\textsuperscript{29} Hallak and Sivadasan (2013) document that exporters have higher average output prices and are more likely to have ISO 9000 certification than non-exporters, even conditioning on plant size. These facts are difficult to reconcile with a model where firm heterogeneity is one-dimensional, as in Kugler and Verhoogen (2012), but fit naturally with a model they develop with heterogeneity in two dimensions: in “process productivity,” which reduces variable costs conditional on quality, and in “product productivity,” which reduces the fixed costs of producing quality (i.e. which, in the notation of Section 2.1, reduces the fixed costs, $f_{ijt}$, required to produce high-quality varieties). In Chinese customs data, Manova and Zhang (2012) show that, within industries, firms that export more and charge higher export prices on average also pay higher prices for their imported inputs, and Manova and Yu (2017) show that, across products within firms, export prices are positively correlated with an index of input prices, constructed using a sector-level input-output table.\textsuperscript{30} Exploiting

\textsuperscript{27}Subsequent papers that have developed heterogeneous-firm models with endogenous output and input quality choice include Kugler and Verhoogen (2012), Hallak and Sivadasan (2013) (discussed below), Johnson (2012), Antonides (2015), Fan et al. (2015), Bastos et al. (2018), and Blaum et al. (2019).

\textsuperscript{28}The within-plant wage change was stronger for white-collar workers than blue-collar workers, hence wage inequality also increased within plants, a finding further explored in employer-employee data in Frias et al. (2012).

\textsuperscript{29}In value-added-tax data from Turkey, Demir et al. (2019) find assortative matching between high-wage buyers and high-wage suppliers, again consistent with the idea that producers of high-quality outputs buy high-quality inputs.

\textsuperscript{30}In Chinese and US data, Bloom et al. (forthcoming) show that many of the relationships previously documented between exports, inputs, and plant size also hold between exports, inputs and measures of management practices, consistent with the idea that larger plants tend to have higher capability than smaller plants, and that higher-capability plants tend to select higher-scoring management practices. Eckel et al. (2015) show that the correlation
barcode-level scanner data from the US, Faber and Fally (2017) find that richer households purchase products from larger firms than poorer households within detailed product categories, again consistent with the quality story. It would be valuable to investigate whether this pattern holds in barcode-level data in poorer countries as well.

An important question in this literature has been whether the upgrading response is attributable to the greater willingness of richer consumers to pay for quality or to two other mechanisms: scale effects, if for instance producing high quality requires paying fixed costs; or distance effects, if for instance per-unit shipping costs are higher for more distant destinations. A small literature has used exchange rate movements as a source of exogenous variation in export destinations to separate these channels. Using panel data on Argentinian firms, Brambilla et al. (2012) show that the Brazilian devaluation of 1999 shifted the composition of export destinations of Argentinian firms toward richer destinations, especially for those firms previously exporting to Brazil. They are thus able to separate the effect of exporting to a richer destination from exporting per se, and they find that the former is associated with an increase in skill intensity and wages while the latter is not. In Portuguese data, Bastos et al. (2018) also use the initial composition of destinations together with exchange-rate movements to show that exporting to richer countries leads countries to pay more for their material inputs, again consistent with a quality story. They find no evidence that exogenous changes in exports per se or in average destination distance lead firms to pay more for inputs. Although firms may charge different markups in different markets, and this may in part explain the output-price patterns, the authors argue that differences in markups alone are unlikely to account for the response of input prices to the export shocks.

The above studies have not had access to direct information on quality and have had to draw indirect inferences from prices and other observables. In the absence of direct information on quality, it is difficult to rule out other explanations for the price patterns definitively. A small but promising literature has had access to direct quality measures, and has corroborated several of the above points. Using wine-guide quality ratings of French champagnes, Crozet et al. (2012) show that firms with higher overall quality ratings charge higher prices, are more likely to export, and export higher volumes and export to more countries. Using wine-guide ratings from Chile, research by Ana Cusolito, Álvaro Garcia-Marin, and Luciana Juvenal, summarized in Cusolito and Maloney (2018), shows that higher-rated wines carry higher prices and are associated with higher material costs. Among soccer-ball producers in Pakistan, where several quality types are directly reported, Atkin et al. (2015) show that, in the cross-section of firms, larger producers produce a higher share of high-quality balls, at a higher average cost, and charge higher prices and markups. Hansman et al. (forthcoming) show that among fishmeal producers in Peru, where protein content is an observable indicator of quality, processing firms are more likely to vertically integrate by buying fishing boats when demand for quality on the export market is high. This integration arguably solves a quality-assurance problem that arises because of imperfect observability of input quality.

Perhaps the cleanest study of the effect of exporting on quality choices is by Atkin et al. (2017a). The authors convinced a US-based non-governmental organization to randomize initial export contracts to Egyptian rug producers and tracked their responses. They paid a local master artisan to evaluate the quality of rugs on a number of dimensions, including the straightness of
corners and how tightly packed the threads were. They find clear increases in product quality and profits among treated firms. They also find effects on productivity, to which we return below.

The idea that demand matters — in particular, that demand from richer end-consumers (at the end of value chains) matters — is reinforced by case studies of Argentinian export industries by Artopoulos et al. (2013), who find that a distinguishing feature of industry pioneers in exporting is that they had direct knowledge of end-consumer tastes in developed-country markets.33 Relatedly, the Enterprise Maps series by John Sutton and co-authors has found that most large firms in several African countries started out as trading firms, rather than as small producers; these findings are consistent with the idea that knowledge of foreign markets is key to firm growth in developing countries (Sutton and Kellow, 2010; Sutton and Kpentej, 2012; Sutton and Olomi, 2012; Sutton and Langmead, 2013; Sutton, 2014).

3.1.1.2 Exports and Technology Adoption There is a small literature on the effect of exports on direct measures of technology and innovation. Bustos (2011) analyzes the behavior of Argentinian firms in response to a regional trade agreement. She first develops a Melitz (2003)-type heterogeneous-firm model in which firms choose between a low-fixed-cost high-variable-cost traditional technology and a high-fixed-cost low-variable-cost modern technology (as previously considered by Yeaple (2005) in a model with perfect competition and ex-ante-homogeneous firms). The theoretical predictions are driven by scale effects: the reduction of tariffs by a trading partner leads exporting firms to expand and to adopt the modern technology. Empirically, Bustos finds that sectors with greater reductions in Brazilian tariffs saw greater increases in exporting, in spending on technology, and in indicators of process and product innovation. Consistent with the theory, these effects are driven primarily by firms in the third quartile of the size distribution (just above the median) in each sector, which in the Argentinian context tend to be the ones that move from non-exporting to exporting. In Canadian data, Lileeva and Trefler (2010) reinforce the basic finding that exports lead to technology adoption. They are able to construct firm-level changes in US tariffs, and find that firms facing greater tariff reductions were more likely to adopt new technologies and to engage in product innovation. They find similar effects on labor productivity, but do not have information on capital stocks with which to estimate TFP. These effects were larger for firms that were initially less productive.34

3.1.1.3 Exports and Productivity In contrast to the literatures on exporting and quality or technology adoption, which consistently find positive effects of exporting, the literature on exports and productivity is mixed, possibly in part because of the measurement issues highlighted in Section 2.2.1.35 As mentioned above, the early literature found little evidence of within-firm effects on productivity (Bernard and Jensen, 1995, 1999; Clerides et al., 1998). More recently, De Loecker (2007) compares Slovenian firms that start exporting to firms that remain only in the domestic market, matching on the propensity to export and controlling for common trends, and finds that the productivity of new exporters rises significantly, especially for firms that start exporting to richer markets. Notably, the paper modifies the Olley and Pakes (1996) procedure by including export status in the construction of the proxy for unobserved productivity in the first stage. (See also De Loecker (2011).) Other papers that have found positive effects of exporting

33See also Sabel et al., eds (2012).
34There is also a small structural literature on exporting and investments in innovation by firms, which is beyond the scope of this review. See e.g. Aw et al. (2011).
35Readers interested in greater detail are referred to the reviews by De Loecker and Goldberg (2014) and Shu and Steinwender (2019).
An important caveat about these papers is that standard TFP measures may reflect markups as well as technical efficiency, as discussed in Section 2.2.1. A recent paper by García-Marín and Voigtländer (2019) addresses this issue. Using detailed plant-product data from Chile, the authors implement a variant of methods developed by De Loecker and Warzynski (2012) and De Loecker et al. (2016) (which in turn builds on insights from Hall (1988)) to estimate markups and marginal costs and investigate how they respond to exporting. Under the assumption that a first-order condition holds for at least one flexible input, the product-level markup can be expressed as the output elasticity with respect to the flexible input divided by expenditures on the input as a share of sales of the corresponding product. Assuming that materials are used across products in the same proportion as in total variable costs, the authors are able to calculate input expenditures as a share of revenues at the product level, using materials as the flexible input. After estimating output elasticities using the method of Ackerb erg et al. (2015) (using single-product firms with a selection correction, following De Loecker et al. (2016)), they calculate product-level markups and use them to recover product-level measures of marginal costs, which they interpret as a measure of productivity. Using this measure and several different estimators, including a propensity-score matching estimator and an instrumental-variables (IV) estimator using tariff changes in export destinations, they find that marginal costs decline by 15-25% for new exporters. Strikingly, when the authors use a standard TFPR measure, they find no effect of exporting; they argue that because the increases in efficiency are passed on to consumers in the form of lower prices, they do not show up in revenues. This study is a notable step forward for the literature. It is also subject to the concern that it depends heavily on the accuracy of the markup estimates derived from the De Loecker and Warzynski (2012) method, which has recently been criticized by Raival (2019) and Traina (2018). In addition, the criticisms of Gandhi et al. (forthcoming) of the Ackerb erg et al. (2015) method of production-function estimation (discussed above) apply here as well (Flynn et al., 2019). However, using the product-level total variable cost and output quantity information, the authors are able to calculate average variable cost at the product level and show that it is highly correlated with the marginal costs they calculate, which provides support for their method.

The most direct evidence of an effect of exporting on productivity is provided by the study by Atkin et al. (2017a) on Egyptian rugmakers, mentioned above. In part for analytical convenience, Verhoogen (2008) models quality upgrading as a shift between lower- and higher-quality goods that a firm already knows how to produce. But Atkin et al. (2017a) argue, convincingly, that the rugmakers learned something in the process of exporting, using two main approaches. In the first, they estimate the effect of treatment on productivity controlling for detailed product attributes and find that it raises TFP. A possible concern, acknowledged by the authors, is that producers choose the product attributes in response to treatment. \footnote{36Conditioning on a set of covariates that respond to treatment breaks the balance on unobservables between treatment and control groups; see e.g. Angrist and Pischke (2009, Section 3.2.3).} This concern does not apply to their second approach, in which they had rugmakers produce identical rugs using the same looms in a laboratory. They find that treated producers make rugs that score more highly on observable quality dimensions but take no less time to produce them. This is already strong evidence for learning. The authors also document an association between messages between the intermediary
and producers about quality issues and improvement on those dimensions. One could raise the question of whether producers gained a pure increase in capability applicable to all types of rugs or learned something specifically about the tastes of foreign buyers. But the constellation of evidence strongly supports the idea that the producers have learned by exporting. This study is a nice example of the advantages of collecting direct information on quality and productivity in a controlled setting (as well as on communications between buyers and producers).

3.1.2 Demand from Local Buyers, Foreign and Domestic

The literature on domestic demand conditions and upgrading has tended to focus on the effects of the presence of multinational corporations (MNCs) in local markets. The entry of foreign firms through foreign direct investment (FDI) is considered by many to be one of the primary drivers of upgrading. But foreign entry may have several effects on local firms. On one hand, foreign entry may generate technological learning spillovers or increased demand (especially for high-quality products) from local firms. On the other hand, foreign firms may have a “business-stealing” effect, gaining market share at the expense of local firms and making it harder for them to reap scale economies.

Early papers using firm-level data found mixed results. In Venezuelan data, Aitken and Harrison (1999) find a negative effect of FDI on the TFP of domestic firms in the same sector, consistent with a business-stealing effect. In Lithuanian data, Javorcik (2004) uses a sector-level input-output matrix to construct measures of exposure to FDI in a firm’s own sector, downstream sectors, and upstream sectors. She finds that firms in sectors that supply the FDI sector experience productivity gains (“backward” spillovers), but that there is little evidence of a productivity effect in the same sector (“horizontal” spillovers) or in sectors that buy from the FDI sector (“forward” spillovers). In a related study in the US, Greenstone et al. (2010) compare counties that win competitions to host large plants, many of them foreign, to counties on the shortlists of candidate locations that lose the competitions. They find that incumbent plants in winning counties see significant TFP increases, and that the spillovers appear to pass through worker-flow and technological links, rather than supplier links. Using the same strategy, Bloom et al. (2019) find spillovers in management practices, but only for firms in sectors with high rates of cross-migration for managers in household data. Abebe et al. (2019) pursue a similar strategy in Ethiopia, comparing TFP outcomes in regions that received foreign investment to regions where firms planned to invest but for bureaucratic reasons were delayed; they find positive effects of nearby FDI on the level of TFP in local firms.

Several papers have examined the effects of the entry of big-box retailers on local suppliers. In a detailed case study of Wal-Mart’s entry into Mexico, Javorcik et al. (2008) argue that there was a heterogeneous effect on local suppliers in the soap and detergent industry: the best suppliers began selling to Wal-Mart and faced pressure to reduce prices but also received input on how to upgrade; weaker suppliers continued to sell through traditional retail channels and just faced increased price competition. Iacovone et al. (2015) develop a dynamic industry-evolution model that captures this effect and find reduced-form evidence consistent with it: in regions with more Wal-Mart stores, and in sectors more likely to be selling to Wal-Mart (e.g. frozen foods), larger plants (presumed to produce products of greater “appeal”) increased sales, R&D spending, wages,

---

37 Javorcik suggested that pressure on local suppliers to raise the quality of goods sold to foreign-owned firms may have been part of the reason for this effect.

38 Atkin et al. (2018) document that foreign retailers in Mexico charge prices that are on average 12% lower than modern domestic retailers, for the same barcode-level product in the same location.
and imported input shares (presumed to be correlated with product quality) relative to smaller plants. In Romania, Javorcik and Li (2013) estimate the effect of the entry of global retail chains on local suppliers, using a summary measure of distance from foreign retailers as a driving variable, and find positive effects on the estimated TFP of affected upstream firms.

An important limitation of the above studies is that until recently it has not been possible to see input-output links at the firm level, and the measures of linkages have had to be constructed using sector-level and/or region-level information. A recent paper by Alfaro-Urena et al. (2019) takes advantage of administrative tax data from Costa Rica, which contains firm-level input-output links. The authors compare firms that start supplying to a multinational corporation (MNC) in Costa Rica to firms that never supply to a MNC and find positive effects on sales to other firms, employment, and standard TFP measures. In a supplemental survey of new MNC suppliers, firms report that the MNCs demand high product quality, which in turn requires using high-quality inputs and changes in hiring, sourcing, and organizational practices.39

A persistent challenge in this literature has been to estimate effects on local firm performance that are not confounded by the effects of demand shocks on markups. A new MNC coming to town can be expected to increase demand for local firms, which may in turn induce local firms to increase markups, which are captured by standard TFP measures. Since the process also often involves quality upgrading, simply estimating TFPQ, if quantity information were available, would not solve the problem. One potential way forward is to use natural experiments to analyze the effect of shocks to domestic demand per se, as opposed to shocks to demand from MNCs. The former are typically not expected to raise product quality, and therefore a comparison between TFPR and TFPQ might be more informative about the role of markups than in settings with larger shocks to the demand for quality. Although not focused on upgrading outcomes, several recent studies examine the effects of arguably as-good-as-random or literally random allocation of government procurement contracts to local firms, for instance in Brazil (Ferraz et al., 2015) and Ecuador (Carrillo et al., 2019). This line of research seems promising.

Another sort of buyer-driven effect arises when customers have preferences directly over the technologies used by firms. One example is provided by Higgins (2019), who shows that when a large Mexican social program (Progresa/Prospera) began disbursing funds on debit cards, corner stores responded by adopting electronic payment technologies, to make payment more convenient for the beneficiaries. (Supermarkets were already largely saturated with the technologies.) Interestingly, the greater use of electronic-payment technologies by corner stores increased demand by other (non-beneficiary) consumers for debit cards, creating a two-sided feedback loop. Another example is provided by the preferences of multinational buyers of consumer goods over working conditions: several studies have found evidence that anti-sweatshop pressure has increased wages and improved working conditions (Harrison and Scorse, 2010; Tanaka, forthcoming).40

### 3.1.3 Competition in output markets

The degree of competition in output markets is another potential driver of upgrading. The key question in this literature, as memorably phrased in the title of Lawrence (2000), is “Does a kick in

---

39In related work in the coffee sector of Colombia, Macchiavello and Miquel-Florensa (2019) show that a quality-upgrading program of a large multinational buyer, which both provided training to farmers and guaranteed a price premium for coffee fulfilling quality (and traceability) requirements, was successfully in increasing the supply of high-quality coffee.

40Relatedly, Boudreau (2019) randomized enforcement of local labor laws by multinational companies in Bangladesh, and found positive effects on compliance with a local requirement to maintain worker-manager safety committees.
the pants get you going or does it just hurt?” The conceptual link between increased competition and upgrading is not obvious. One common argument is that firms do not maximize profits prior to the increase in competition and are spurred to do so (to increase “X-efficiency” in the terminology of Leibenstein (1966)) by the competitive threat. But this argument also needs to explain why firms were not maximizing profits in the first place. One also needs a mechanism strong enough to overcome the possible reduction in scale — and hence in scale economies — by firms facing stronger competition. Empirically, the challenge is to separate the effect of competitive pressure to upgrade from the effect of killing off firms that fail to upgrade. Holmes and Schmitz (2010) review the theoretical and empirical research on these issues, focused mainly on developed countries. Although they discuss a number of ideas, they acknowledge that there is little consensus in the literature about theoretical mechanisms.

Empirically, there is reasonably convincing evidence of a positive effect of competition on firm performance in particular cases. One leading study is Schmitz (2005), which tracks the response of US iron ore firms to the lower prices of Brazilian ore in the 1980s. Schmitz finds significant increases in productivity and argues that they were mainly due to changes in work practices, made possible in part because the competitive threat led unions to be more flexible about work rules. He marshals direct evidence from collective bargaining contracts and staffing levels, which reinforces the findings from more conventional productivity estimation. In a developing-country context, Das et al. (2013) focus on a public-sector rail mill in India which was for many years the exclusive producer of long rails for Indian railroads. In the late 1990s, the Indian government invited private companies to begin production and a large private conglomerate announced its intention to enter. Output per shift in the rail plant, measured in physical units, rose by 30% in a matter of months. Another example is provided by Jensen and Miller (2018), who study boat-builders in Kerala, India. The expansion of cellphone coverage led fishermen to travel further so they could sell their fish at the best prices. This increased their knowledge of boat-builders in other villages and arguably increased competition in the boat-building market. In turn, increased competition led to an expansion of the businesses of higher-skilled (and higher-quality) boat-builders and a contraction of those of lower skill, raising average quality. The greater scale for higher-skill builders also arguably enabled greater capacity utilization and greater labor specialization within firms, reducing costs. Another interesting example, from the Chinese footwear industry, is offered by Qian (2008). Following a shift of intellectual property rights enforcement resources away from counterfeiting in 1995, the industry saw a sharp increase in the entry of low-quality producers selling counterfeit brands. To differentiate themselves, more-productive, higher-quality producers upgraded quality and vertically integrated downstream by opening company stores.

A large number of papers have explored the consequences of reductions of import tariffs on within-firm productivity changes. These studies have typically considered many sectors together, and do not have the sort of detailed information on business practices or physical output that the papers discussed above have. An early paper by Pavcnik (2002) used the Olley and Pakes (1996) methodology to estimate TFP in Chilean data and found that productivity increased in import-competing industries relative to non-traded industries following Chile’s unilateral liberalization in

---

41 This could be either because they fail to optimize altogether or that they optimize an objective other than profits. This issue is discussed further in Section 3.3.1 below.

42 Using case studies of the construction equipment, automotive, and machine tools industries in China, Brandt and Thun (2010) develop the related and interesting idea that competition at the low-quality end of industries induced domestic firms to upgrade to the middle-quality segment to escape competition. The fact that China has a large domestic market meant that firms were shielded somewhat from foreign competition even in the middle-quality segment, because foreign firms had higher costs and less knowledge of domestic consumers. See also Medina (2018).
the late 1970s. (See also Tybout et al. (1991).) Amiti and Konings (2007), in one of the first papers to separate the effects of tariffs on a firm’s outputs and inputs, apply the Olley and Pakes (1996) methodology to estimate TFP in Indonesian data and estimate separately the effects of tariffs on outputs and inputs. The effects of output-tariff reductions on productivity are positive but modest, especially relative to the input-tariff effects (mentioned in Section 3.2.1 below). Papers that have found a positive effect of output-tariff reductions on productivity include Schor (2004) and Muendler (2004) in Brazil, Fernandes (2007) in Chile, Lopez Cordova (2003) and Iacovone (2012) in Mexico, Yu (2015) in China, and Topalova and Khandelwal (2011), Nataraj (2011), and De Loecker et al. (2016) in India. A small literature has also found effects of output tariff reductions on R&D expenditures and/or other innovation outcomes in developing and emerging countries (Teshima, 2010; Gorodnichenko et al., 2010).

But there is reason for caution in concluding that trade competition has an unambiguously positive effect on productivity. In the corrected version of the study of WTO accession on Chinese firms by Brandt et al. (2017, 2019), the effect of output-tariff reductions on the productivity of incumbent firms is not statistically distinguishable from zero. In detailed Ecuadorian data, Bas and Paunov (2019) find mostly statistically insignificant results of output tariffs on TFP measures. Holmes and Schmitz (2010) note that studies often focus on tariff effects on productivity changes in surviving firms, which may be a selected sample. The extent to which import competition raises productivity by killing off less-capable firms versus stimulating firms to improve their performance remains a persistent question. The issues with standard TFP measures discussed in Section 2.2.1 continue to be concerns in many studies. There is also well-identified historical evidence that temporary protection from British imports during the Napoleonic wars promoted adoption of mechanized cotton spinning in Northern France (Juhász, 2018), suggesting that reduction of competition can also increase productivity.

Overall, the evidence on the effects of competition on upgrading seems somewhat inconclusive. It is clear that increased competition can have positive effects on firm performance in some cases, but the effects vary significantly across settings. More research is needed to better understand the conditions under which competition stimulates upgrading. One interesting idea, which has not been well explored empirically at the firm level, is that competition plays more of a stimulating role for firms closer to the world technological frontier than for those further away (Aghion et al., 2005a,b; Amiti and Khandelwal, 2013).

3.1.4 Reputation in Output Markets

The quality models discussed above treat quality as observable and enforceable in contracts. But in the real world, information is often asymmetric. Buyers may only learn about the quality of a good after a transaction has taken place, and, if the quality is lower than contracted, may have difficulties getting a court to enforce the contract. The same goes for other product characteristics (broadly construed) such as the timeliness of delivery. These issues are especially severe in developing countries, where quality and reliability vary greatly across firms and legal institutions are weak.

In such settings, firms typically rely on repeated interactions and the threat of discontinuing a

---

43In a similar spirit, Bloom et al. (2016a) find positive effects of competition from China on patenting, information technology use, and TFP in twelve European countries. In Spanish data, Chen and Steinwender (2019) find positive effects of import competition on productivity for initially less-productive, family-managed firms. By contrast, Autor et al. (forthcoming) find negative effects of Chinese competition on patenting in the US.

44See also Yang et al. (2019).

45For a useful overview of the international dimensions of these contracting issues, see Anträ (2015).
relationship to enforce agreements; in other words, they enter into relational contracts (MacLeod and Malcolmson, 1989; Baker et al., 2002). But establishing a relational contract, and developing a reputation for quality and reliability, can take time and require up-front investments. This can be especially challenging in developing countries, because buyers often use average quality in a country or country-sector to form expectations about the quality of a particular firm. Given this collective-reputation issue, it may not optimal for individual firms to upgrade: there may be a low-quality equilibrium trap (Tirole, 1996). In such situations, mechanisms that allow firms to build individual reputations may stimulate upgrading. In addition, networks of firms may facilitate contracting, by providing information about potential trading partners, enhancing a firm’s ability to sanction partners who renege, and giving the group an incentive to sanction its own members in order to maintain a group reputation.

A small but growing literature has explored these issues empirically in developing countries. Using a tailored survey of Vietnamese firms, McMillan and Woodruff (1999) document that, consistent with models of relational contracts, firms’ willingness to supply trade credit (an indicator of how much the firm trusts a trading partner) depends on a number of features of the relationship: how easy it is for the partner to find another supplier, how long the two parties have been transacting, and the density of network links. In data on contracts of Indian software firms, Banerjee and Duflo (2000) show that older firms and firms with a very long-term, open-ended relationship with the buyer — characteristics plausibly associated with the reputation of the Indian firm — are offered more attractive contracts, in the sense that the buyer is more willing to accept responsibility for cost overruns. Macchiavello (2010) shows that Chilean wineries receive more attractive terms from UK wine distributors over time, controlling for such factors as quality and winery-distributor match effects, suggesting that the wineries acquire improved reputations over time.

Macchiavello and Morjaria (2015) examine the response of Kenyan rose exporters to a major supply disruption brought about by ethnic violence in 2008 and find patterns consistent with a reputation model. In particular, they find an inverted-U relationship between relationship age and the exporters’ compliance with agreements to provide flowers during the violence (which raised the cost of supplying flowers). Compliance initially increases with age because the value of the relationship increases with age. But at a certain point, sellers have established their reputations with the buyers, and do not have to worry as much about damaging their reputation by not complying.46

A recent experiment by Bai (2018) with watermelon sellers in China highlights the importance of branding for the development of reputations: simply giving sellers a hard-to-counterfeit way of marking their watermelons was sufficient to induce them to upgrade the quality of goods sold with that mark. A somewhat contrasting case is offered by Bold et al. (2017), who calibrate a learning model using data from agricultural trials in Uganda and argue that, given the noise in the environment and the difficulties that consumers have in inferring fertilizer quality, it would be very costly for a seller of fertilizer to develop a reputation for supplying high quality. This may explain the fact that the fertilizer market appears to be stuck in a low-quality equilibrium. Bai et al. (2017) provide evidence for the role of group reputation in the Chinese dairy industry. In 2008, a subset of producers were found to have sold adulterated baby formula by adding the industrial chemical melamine. Exports dropped by 68% following the scandal, and, perhaps surprisingly, firms that were inspected by the Chinese authorities and found to be innocent saw similar declines as those found to be guilty. The group reputation effects appear to have been particularly strong.

46In related work, Ghani and Reed (2019) examine how relational contracts between ice sellers and fishermen in Sierra Leone evolve in response to an increase in upstream supply of ice.
in this case.

Overall, despite these notable contributions, we are still at an early stage of learning about the causal mechanisms linking the costs of acquiring a reputation in output markets and upgrading by industrial firms. Newly available data from online platforms are making it possible to investigate reputation mechanisms at a level of detail not previously possible; see Tadelis (2016) for a review. This area seems to be very fertile ground for research.

3.2 Input-Side Drivers

We turn now to drivers on the input side, beginning with factors influencing imports of inputs and then considering factors that influence the prices and availability of domestic inputs.

3.2.1 Imported Inputs

Above we observed that firms in developing countries appear on average to sell higher-quality varieties on international markets than on domestic markets. It also appears that firms tend to buy higher-quality inputs on international markets than on domestic ones. In Colombian data, for instance, Kugler and Verhoogen (2009) document that plants systematically pay higher prices for imported inputs, controlling for detailed product fixed effects. One possible explanation is that there is a home-market effect in the production of quality, such that firms in richer countries specialize in producing higher-quality goods to appeal to richer local consumers, as for instance modeled in Fajgelbaum et al. (2011).

If inputs available on the international market tend to be of higher quality than domestic inputs (or have lower quality-adjusted prices for higher-quality varieties), then we would expect a reduction of tariffs on inputs to lead developing-country firms to upgrade the quality of their inputs. Given the fact (discussed above) that high-quality inputs appear to be a key requirement for high-quality outputs, we would expect to see upgrading on the output side as well. Bas and Strauss-Kahn (2015) provide evidence for this mechanism in Chinese trade-transactions data. Comparing processing firms (which are exempt from tariffs) to ordinary firms, constructing firm-specific tariff reductions based on firms’ import mixes, and controlling for firm-product (and a number of other) fixed effects, they find that tariff reductions lead Chinese firms to increase the prices they pay for inputs and to increase the prices they charge for outputs, consistent with a quality story. The results are primarily driven by firms that import most of their inputs from, and export most of their outputs to, developed countries. Results are similar if they use the Khandelwal et al. (2013) methodology to construct measures of input and output quality. A roughly contemporaneous paper by Fan et al. (2015) also finds that Chinese firms responded to reduced tariffs on imported inputs by raising export prices and quality, and that this effect is stronger in more differentiated sectors. (See also Feng et al. (2016) and Abeberese (2016).) An obvious limitation of trade-transactions data is that they include only international transactions, which may not be representative. However, Bas and Paunov (2019) find broadly similar results with representative data from Ecuador (plant census and customs data), and also find that the imported-input-driven upgrading is associated with increases in skill intensity.

In an interesting extension of this line of work, Fieler et al. (2018) argue that there is an amplification effect in upgrading: tariff reductions on inputs lead firms to upgrade the quality of

\[47^{\text{Importing plants also pay more on average for their inputs than non-importing plants, even for domestic inputs, consistent with the ideas that there are fixed costs of importing and that more-capable plants use imported inputs, which tend to be higher-quality, to produce higher-quality products. See also Blaum et al. (2019).}}\]
outputs, which in turn increases their demand for other high-quality inputs, which gives incentives for local suppliers to upgrade, which gives local final-good producers further incentives to upgrade. Empirically, the authors calibrate their model to pre-liberalization data and do counterfactual simulations. Now that datasets with firm-to-firm links are becoming available, a promising line of research would be to investigate this sort of mechanism in a less theory-dependent way.

Tariff reductions not only improve access to high-quality imported inputs, they also expand the variety of inputs available, which may in turn enable firms to produce new outputs. Focusing on India’s liberalization in the early 1990s, Goldberg et al. (2010) provide evidence that the increased availability of imported inputs led firms to expand their set of output varieties. They document a reduced-form relationship between import tariff reductions and product innovation and impose a simple theoretical structure to separate the price and variety effects of the tariff reductions, finding that a substantial share of the increase in product scope is driven by the expansion of imported input variety. Although Goldberg et al. (2010) do not have access to information on inputs at the firm level, Bas and Paunov (2019) directly observe both inputs and outputs of Ecuadorian firms and confirm the findings that import tariff reductions lead firms to use more inputs and expand product scope.

There also appears to be a robust causal relationship at the firm level between reductions of tariffs on imported inputs and increases in standard measures of revenue TFP. This relationship has been documented for instance by Schor (2004) in Brazil, Amiti and Konings (2007) in Indonesia, Topalova and Khandelwal (2011) and Nataraj (2011) in India, Brandt et al. (2017, 2019) in China, and Bas and Paunov (2019) in Ecuador.48 In a recent review, Shu and Steinwender (2019) observe that papers that have considered tariffs on outputs and inputs separately have tended to find stronger effects of input-tariff reductions than of output-tariff reductions, and I share their view. (Refer to Section 3.1.3 above.)

At the same time, a recurrent question in the literature on imports and productivity is to what extent the results reflect changes in markups or some other source of bias in measured TFP, for instance quality. In an influential contribution, De Loecker et al. (2016) develop a methodology to tease apart the contributions of technical efficiency, markups, and quality in multi-product firms. At the core of the exercise is a formula for calculating markups at the firm-product level, discussed in the context of Garcia-Marin and Voigtländer (2019) in Section 3.1.1.3 above. The formula requires information on input expenditures as a share of output revenues and on output elasticities at the product level. The authors’ strategy is to focus on single-product firms, where the mapping from inputs to outputs is clear, and to do a selection correction to address the fact that single-product firms may not be representative. In the output-elasticity estimation, which follows Ackerberg et al. (2015), the authors put physical output on the left-hand side and use output prices to proxy for input prices and input quality to address potential input-quality bias. They find that import tariff reductions cause a reduction of marginal cost that is only partially passed through to consumers. That is, product prices decline, but by less than marginal costs decline, and hence markups rise. This suggests that the estimated effects of import tariff reductions on standard TFP measures — which incorporate both technical efficiency and markups — overstate the true effect on technical efficiency. Quibbles can be raised about the output-elasticity estimation (which is subject to the identification concerns raised by Gandhi et al. (forthcoming)) and about whether putting physical output on the left-hand side in the production-function estimation adequately

48 See also the studies by Tybout and Westbrook (1995), Lopez Cordova (2003), Kasahara and Rodrigue (2008) and Halpern et al. (2015), which find positive contributions of imported inputs to productivity. An exception is Muendler (2004), which finds that imported inputs make only a minor contribution to productivity, if any.
addresses the possibility of output-quality bias. But it is clear that this paper is an important contribution and has become a key point of reference for the literature.

### 3.2.2 Domestic Inputs

Several papers have investigated how changes in the cost of labor, capital, or other inputs on the domestic market affect firms’ upgrading decisions. Supply shocks of workers of different skill levels are one possible driver. Some of the best work on this topic is from the US: using a shift-share instrument for immigration, Lewis (2011) shows that US manufacturing firms in regions with greater inflows of low-skilled migrants were less likely to adopt advanced technologies, and Hornbeck and Naidu (2014) show that greater outflows of low-skilled workers from the US South, in response to a major flood in 1927, led farms to increase mechanization. In a similar vein in a developing-country context, Imbert et al. (2019) use agricultural price shocks combined with historical migration patterns in China as a source of exogenous inflows of low-skilled migrants to urban areas. Firms in areas that receive more low-skilled migrants are less likely to file domestic patents and tend to shift toward products with low human-capital intensity (defined as the average share of the workforce with a high-school degree among firms that produce a given product).

Two recent papers using city-level minimum-wage variation in China provide evidence that minimum wage regulations, which raise the relative cost of less-skilled labor (in addition to raising wage costs overall), can have effects similar to an increase in relative supply of more-skilled labor. Mayneris et al. (2018) find that firms more exposed to the minimum-wage hikes (in particular, those whose average wage in the previous year was below the new minimum wage) saw increases in productivity relative to less-exposed firms. Hau et al. (forthcoming) also find that firms more affected by minimum wage changes (in the sense that their average wages are closer to the minimum) tended to see increases in measured TFP and shifted to more capital-intensive production, with some heterogeneity based on firm characteristics. The usual caveats about TFP estimation apply, but broadly these papers suggest that higher wages overall (which induce firms to substitute capital for labor) and/or higher relative costs of low-skilled workers (which induce firms to substitute high-skilled for low-skilled labor) can lead firms to upgrade.

The literature on access to capital as a driver of upgrading in larger firms in developing countries remains thin and somewhat mixed. There have been influential studies of the effect of capital in microenterprises. There have also been careful studies of the effects of capital-supply shocks on other (i.e., non-upgrading) outcomes among larger firms, in both developed and developing countries (e.g., output: Banerjee and Duflo (2014); use of alternative credit sources and financial distress: Khwaja and Mian (2008); exports: Amiti and Weinstein (2011), Zia (2008), Paravisini et al. (2014), Kapoor et al. (2017); employment: Chodorow-Reich (2014), Brown and Earle (2017)). But there have been relatively few studies linking credit shocks directly to firm-level productivity, quality, technology adoption, or other upgrading outcomes among larger developing-country manufacturing firms.

---

49 See also Clemens et al. (2018) and San (2020).

50 Related work by Bustos et al. (2019), with data at a regional level in Brazil, suggests that such shifts into low-skill-intensive manufacturing may have lock-in effects with negative growth consequences in the long run.

51 To be clear, although higher minimum wages appear to have spurred upgrading in these cases, they are likely to have reduced profits for individual firms. The point from Section 2.1 that upgrading may or may not be profit-maximizing is worth recalling here.

52 See e.g. de Mel et al. (2008), McKenzie (2017), and the reviews by Banerjee et al. (2015), Woodruff (2018) and Quinn and Woodruff (forthcoming).

53 There are small literatures on credit constraints and technology adoption in agriculture (see e.g. Giné and
Perhaps surprisingly, the few papers that have focused on the effect of increased capital supply on productivity have largely failed to find evidence of such an effect. Bau and Matray (2019) examine the effect of a policy reform in India that removed some restrictions on foreign investment, arguably increasing the supply of capital, in a staggered way across industries. They primarily focus on misallocation, but they also estimate the impact of the reform on TFPR, and find no evidence of an effect. They caution that they also find a decline in product prices, likely reflecting reduced capital costs, and that the price decline may in part be responsible for the lack of an observed effect on TFPR. Also in India, Rotemberg (forthcoming) examines the effects of a 2006 broadening of the set of firms in India eligible for subsidies to small and medium-sized businesses, similar to an earlier change studied by Banerjee and Duflo (2014). The affected firms became eligible for a range of programs, but the most important (70% of the budget for such programs) appears to have been subsidized credit. Rotemberg focuses primarily on quantifying simultaneously the direct and indirect effects of the subsidies and their contributions to aggregate productivity, but he also examines direct effects of the subsidies on firm-level TFPQ and finds no evidence of an effect. Cai and Harrison (forthcoming) study a reform in China that reduced the value-added tax (VAT) on investment goods, with the goal of encouraging technology adoption. They find an increase in capital intensity but no effects on fixed investment, product introductions, or productivity.54 Arráz et al. (2014) study the effect of a Colombian government loan-guarantee fund, using a propensity-score matching estimator with fixed effects, and find impacts on output and employment but not investment, productivity, or wages. By contrast, Eslava et al. (2012), also using a combination of matching techniques and fixed-effect estimators, find that loans from a publicly owned development bank to Colombian manufacturing firms generated significant positive effects on productivity as well as output, employment, and investment.

Energy inputs are often measured reasonably well in manufacturing surveys in developing countries, and a small literature has investigated the role of shocks to energy supply or prices on firm-level upgrading outcomes. Abeberese (2012, 2017) examines the relationship between electricity prices and various dimensions of firm behavior, using arguably exogenous variation in coal prices interacted with the initial share of thermal generation (which uses coal) in states’ electricity generation. She finds that higher electricity prices induce firms to shift their product mix toward products that are on average produced by firms that use less electricity. Although specific technologies are not observed in the Indian data, it seems plausible that less electricity-intensive processes are also less technologically advanced. She also finds a negative (although not significant) relationship between electricity prices and the level of productivity, and a negative and significant relationship between electricity prices and the growth rate of productivity.55 A subsequent paper by Allcott et al. (2016) pursues a related strategy. Using rainfall at higher elevations (which determines hydro-electric power generation capacity) as an instrument for shortages (rather than electricity prices) in India, they find that shortages lead firms to contract in terms of both sales and input purchases but they do not find a significant effect on TFPR. Simulations suggest that there is more of a negative effect for firms that do not already have generators, which are smaller on average.56

Klonner (2008) and the review in Jack (2013, Section 5)) and households (see e.g. Berkouwer and Dean (2019)).


55In related work in Chinese data, Fisher-Vanden et al. (2015) find that firms respond to higher electricity prices by outsourcing more inputs; at the same time, they find muted effects on productivity. Related contributions not focused on firm-level upgrading outcomes include Rud (2012) and Cole et al. (2018).

56Relatedly, Abeberese et al. (forthcoming) find negative impacts of outages on productivity among small and medium-sized Ghanaian firms (see also Hardy and McCasland (forthcoming), which focuses on microenterprises)
3.3 Drivers of Firm Capabilities

This section reviews research on factors that operate through effects on firm capabilities and knowledge. A first issue that arises is the motivation of entrepreneurs, in particular whether or not they can be presumed to maximize profits. We then turn to various factors that influence firms’ know-how.

3.3.1 Objectives of Entrepreneurs

The framework in Section 2.1 assumes that firm seeks to maximize the discounted present value of profits, expressed in equation (2). Is this a plausible assumption? One reason it may not be is that entrepreneurs consciously hold other objectives. Entrepreneurs may value a quiet life (Bertrand and Mullainathan, 2003) or derive private benefits from control or empire-building (Williamson, 1964). Although these motivations are often attributed to non-owner managers, they might also characterize owners themselves. Another possible reason is that entrepreneurs would like to maximize profits but have behavioral biases that lead them to make mistakes. While these possibilities are widely acknowledged, there is relatively little empirical research directly on the question of whether individual owners of medium-sized or large firms hold non-profit-maximizing objectives or systematically make mistakes.\(^5\) There is evidence suggesting that mistakes are made by small shopkeepers, in the form of lost sales due to holding insufficient change (Beaman et al., 2014), and by agricultural producers, in the sense of failing to notice relevant information about production (Hanna et al., 2014) or failing (because of time-inconsistent preferences) to invest in fertilizer (Duflo et al., 2011). But more empirical investigation of the objectives consciously held by firm-owners and of their behavioral biases is sorely needed.

Two words of caution are in order. First, the question of whether an individual entrepreneur maximizes utility is distinct from the question of whether a firm profit-maximizes. As we will see below, a firm may fail to take advantage of an apparent profit-making opportunity, even if all individuals within the firm are behaving rationally, in pursuit of standard objectives. Second, it appears to have become more common in recent years to attribute poor firm performance in developing countries to failures of entrepreneurs to profit-maximize. But as noted above, entrepreneurs in developing countries often face very different conditions in product and input markets, and hold different amounts of know-how, from rich-country entrepreneurs. We need to examine very closely the constraints they face before we can conclude that they have failed to optimize. In an agricultural context, Schultz (1964), Stiglitz (1989) and others have argued for a “poor but rational” view: if we observe behavior that seems to be non-optimal, we should ask ourselves what problem is being solved, and what constraints producers face, before concluding that they are not optimizing. A similar point applies to entrepreneurs in larger manufacturing firms. This is not to say that all developing-country entrepreneurs are perfect exemplars of Homo Economicus, but rather that we should be cautious before concluding that they are not.

---

\(^{5}\)The recent review by Kremer et al. (2019) devotes a section to “behavioral firms” but asserts that “we have a limited understanding of what the objectives of firm-owners in developing countries are” (p. 418).
3.3.2 Entrepreneurial Ability

Turning to drivers of capabilities, a first one to consider is entrepreneurial ability, which we can think of as a fixed characteristic of an individual entrepreneur — in the framework of Section 2.1, a time-invariant component of capability that is common across products and techniques. Recent research has taken several approaches to evaluating the importance of entrepreneurial ability. One approach is to examine cross-sectional correlations between detailed manager characteristics and firm performance. For instance, there is evidence from a range of countries, including Brazil and India, that firm performance is positively correlated with the amount of time CEOs spend in high-level meetings, rather than production activities (Bandiera et al., forthcoming). Focusing on six factories of an Indian garment firm, Adhvaryu et al. (2019a) find that factor-analytic summary measures they characterize as managerial attentiveness and autonomy correlate positively with levels of productivity and the rate of productivity improvement on production lines.58 A natural question that remains open is whether the correlations reflect causal effects of manager characteristics or some form of sorting of managers to firms or production lines.

Another way to assess the role of such fixed manager characteristics is to examine changes in firm decisions and performance in response to changes in top managers. This is the strategy of Bertrand and Schoar (2003), who find in US data that manager fixed effects have significant explanatory power for various corporate decisions, even controlling for rich sets of firm observables. A small literature examines the decisions and performance of family-owned firms where managerial positions are passed between family members (as opposed to being filled through competitive searches). There is robust evidence that inherited control is bad for performance (Pérez-González, 2006; Bennedsen et al., 2007; Bertrand et al., 2008). There is also evidence that family control is associated with lower scores on the World Management Survey index (Bloom and Van Reenen, 2007, 2010; Bandiera et al., 2017). Instrumenting family control with the gender mix of the previous CEOs’ children, Lemos and Seur (2019) have recently shown that this relationship is causal: family control leads to lower-scoring management practices.

Another type of evidence comes from changes of ownership. Using detailed data on ownership and physical inputs and outputs in the Japanese cotton spinning industry in the Meiji era, Braguinsky et al. (2015) find that acquisitions are associated with increases in TFPQ in the acquired firms. Interestingly, the acquiring firms typically do not have higher physical productivity than the acquired firm prior to purchase, but they are more profitable, in part, the authors suggest, because they are able to manage demand fluctuations to maintain higher levels of capital utilization. Using a propensity-score matching estimator in Spanish data, Guadalupe et al. (2012) find that acquisition by a multinational firm leads to upgrading on a number of directly observable dimensions, including indicators for process and product innovations, purchases of new machinery, and the introduction of new organizational practices. Studies in developing countries have largely found positive effects of foreign ownership on productivity (Arnold and Javorcik, 2009; Javorcik and Poelhekke, 2017; Stiebale and Vencappa, 2018), although there is still a debate about whether acquisition by multinationals has larger impacts than acquisition by domestic firms (Wang and Wang, 2015). In Indian data, Stiebale and Vencappa (2018) also find evidence of a positive effect of foreign acquisition on quality upgrading, indicated both by an increase in input prices and by a measure of product quality along the lines of Khandelwal et al. (2013).

Overall, the evidence seems strong that entrepreneurial ability matters for upgrading outcomes and that family control is associated with worse performance. This raises a question of why family

---

58Relatedly, Adhvaryu et al. (2019b) find that more attentive managers are more effective in reallocating workers in response to negative worker-level productivity shocks from pollution exposure.
control is so prevalent, a topic to which we return in the next subsection.

3.3.3 Agency Issues

Firms are collections of people with sometimes aligned but sometimes conflicting interests. Even if an entrepreneur is rational and of high ability, she may still have difficulties in getting employees to act in a desired way. These agency issues can be thought of as influencing a firm’s capabilities. The extent to which a firm is able to resolve them will clearly matter for its ability to upgrade. The agency literature is very large; here we focus on empirical studies in developing countries on how agency issues influence upgrading outcomes at the firm level.

The Atkin et al. (2017b) study of Pakistani soccer-ball producers highlights the importance of such agency issues. Through a series of fortuitous events, the research team came up with a new technology — a design for cutting more pentagons from a rectangular sheet and a piece of equipment, an “offset” die, to implement that design. An advantage of the context is that all firms use the same, simple production process, at least for part of their production, and it is possible to calculate directly the benefits of adoption, which are positive on net for essentially all firms. The researchers gave out the technology to 35 firms, expecting the treated firms to adopt quickly and planning to track the channels of spillovers. But 15 months later, only 5 treated firms and 1 control firm had adopted, despite the fact that the technology appeared to be working as expected. Conversations with firm owners and employees revealed the reason: the key employees, cutters, were paid piece rates based on the number of pentagons cut, with no incentive to reduce waste, and the offset die slowed them down initially. Although the reductions of waste were much larger than the increases in labor costs, under the existing contracts the cutters’ incomes would have declined with adoption and so they found various ways to discourage it. The researchers conducted a second experiment in which employees received a bonus of a month’s salary if they demonstrated the productivity benefits of the new die in the presence of their employer. The second experiment generated a statistically significant increase in adoption by firms, suggesting that a conflict of interest within the firm had been at least in part responsible for the initially slow adoption of the offset dies. A natural question is why firm owners did not adjust their payment schemes to reward the employees for adopting the new technology (or at least keep them whole). One possibility is that owners simply did not understand the problem; another, consistent with qualitative evidence, is that they understood, but that there are costs to changing employment contracts, even informal ones, and that owners calculated (perhaps with low priors about the value of the technology) that the expected benefits did not compensate for the re-contracting costs. The failure to adopt the new dies is arguably an example of what Garicano and Rayo (2016) call an “organization failure” — the firm as a whole failed to adopt a more-efficient technology — even though all individuals in the firm appear to have been acting rationally, given their knowledge. The case is also arguably an example where contracts that were optimal in a technologically static environment (here, piece rates before the new die) were not optimal in a technologically dynamic one (once the new die was introduced), and the stickiness of contracts generated a sort of organizational inertia.

A recent study of the adoption of credit scoring by Indian banks by Mishra et al. (2019)

---

59 See e.g. the reviews by Gibbons (2010), Gibbons and Henderson (2013), Lazear and Oyer (2013), and Garicano and Rayo (2016). Bandiera et al. (2011) review related work on how social connections and incentives can affect productivity.

60 The cost reduction is modest, approximately 1% of total costs, but the fixed costs of adoption are also modest. The authors calculate the time required to recoup the fixed costs to be less than 8 weeks for 75% of firms in the treatment group.
provides additional evidence for organizational inertia. The key finding is that older banks, both public and private, founded prior to the beginning of India’s liberalization in 1991, are less likely to adopt credit scoring for existing clients than the same banks are for new clients or than new banks (founded post-1991) are for existing clients. The authors suggest that the older banks developed an organizational culture and way of dealing with existing clients under the less competitive pre-liberalization regime and that the culture has persisted, fading away only slowly.

A recent study by de Rochambeau (2017) identifies another sort of agency issue. The author randomly gave out GPS monitors to trucking firms in Liberia. She finds that they reduced unauthorized breaks and average travel times for the trucks on which they were installed, as expected. But she also finds that owners were less likely to install the monitors on trucks of drivers who had better performance at baseline, who tended to come from the same county as the owners (an analogue of co-ethnicity in the Liberian context). For high-initial-performance drivers who received the monitors, their performance on non-monitored tasks deteriorated. It appears that the monitoring had adverse effects on the performance of drivers who were otherwise intrinsically motivated. Owners plausibly sought to avoid such adverse effects by not installing them for many drivers from the same county.

Ethnic divisions within firms appear to matter for performance in other ways as well. Hjort (2014) looks at how the ethnic composition of teams affects output in a flower firm in Kenya. Ethnically homogeneous teams are more productive than heterogeneous ones, and this tendency is exacerbated during a period of ethnic strife in Kenya. The impact on firm productivity is substantial. Hjort argues that the patterns are consistent with a model of taste-based discrimination against non-co-ethnics. The extent to which firms are able to mitigate such conflicts can be thought of as a component of firm capability.

Macchiavello et al. (2015) make a related point regarding gender in the context of an experiment in Bangladeshi garment factories, where most line workers are female and most supervisors are male. Both male and female employees believe, incorrectly, that female supervisors have less technical knowledge. This incorrect belief fades with exposure to female supervisors (who are randomized across production lines in their experiment). But there is a cost of overcoming the prejudices of employees, and it is not clear that it is profit-maximizing for an individual firm to pay the cost of shifting the norm.

Returning to the question of why family ownership is so prevalent, a number of authors have argued that family control is in part a response to agency issues within firms, in particular to the problems that owners may have in inducing the behavior they desire from non-family managers. Ilias (2006) focuses on the surgical goods industry in Sialkot, Pakistan, and argues that the tendency of non-family managers to move to other firms and take clients and production knowledge with them leads families to favor family members as managers. One symptom of this behavior is that founders of firms who have more brothers end up with larger firms."61 Cai et al. (2013) present evidence from Chinese firms that family members who are managers are paid more but have lower-powered incentives than non-family-member managers, consistent with the idea that family members are trusted more to act in the interests of the firm. These findings do not contradict the findings above that continued family control after the founder dies is bad for performance. But they do suggest that there is a reason why family control persists. Like piece rates in the soccer-ball example, family control may be another instance of a solution to agency problems that is initially beneficial (in the sense of reducing malfeasance under the founder) but that outlives its usefulness (once the founder dies).

"61 Bloom et al. (2013) make a similar observation about the Indian textile firms they study.
3.3.4 Learning

For a given level of entrepreneurial ability and degree of resolution of agency problems, a firm’s accumulation of know-how — learning — can drive upgrading. But in many cases, know-how cannot simply be purchased on an open market or downloaded from the internet. Much of the knowledge needed to produce successfully is tacit (i.e. not written down in instruction manuals) an idea that goes back at least to Katz (1984) and Pack and Westphal (1986). In addition, many organizational capabilities need to be worked out in the practice of producing; as Gibbons (2010) puts it, they need to be “homegrown.” (See also Dessein and Prat (2019).) Learning is likely to require investments with uncertain payoffs, and to take time. This subsection reviews recent work on a number of channels through which learning can occur.

3.3.4.1 Learning within firms

An important distinction in the learning literature is between learning from one’s own experiences (i.e. learning by doing) and learning from others. There is extensive evidence from industrialized countries that firms learn by doing and that the rate of learning can vary widely across firms (see e.g. Argote and Epple (1990), Irwin and Klenow (1994), Benkard (2000), Thompson (2001), Levitt et al. (2013), and Hendel and Spiegel (2014)). To date, there has been relatively little research on specific mechanisms of learning-by-doing within larger manufacturing firms in developing countries. One exception is the recent study by Menzel (2019), which uses detailed production data from three multi-floor garment factories in Bangladesh and finds that knowledge about how to produce new designs spills over across production lines on the same floor (which correspond to organizational subdivisions of the companies), but not across floors. Atkin et al. (2017b), discussed above, also documented a form of learning within firms.

Another form of within-firm learning is the transfer of knowledge or technologies across establishments (or across firms within a corporate group). These transfers are easier to observe when they cross international borders. Using data on foreign affiliates of US multinational firms in a large set of countries (including many developing countries), Branstetter et al. (2006) show that when countries strengthen their intellectual property protections, royalty payments for technology transferred to affiliates in those countries increase. There is also evidence for technology transfers across firms within developing countries. For instance, Jiang et al. (2018) look at innovation outcomes in international joint ventures in China, and also for firms that participate in the joint ventures (separate from the joint venture themselves), and find that such partner firms see within-firm increases in patenting rates following the establishment of the joint venture. (See also Bai et al. (2019).)

3.3.4.2 Learning from other firms

Besides learning from their own experiences, firms also clearly learn from others. Although perhaps the strongest evidence of such learning spillovers comes from developed countries (Irwin and Klenow, 1994) or agriculture in developing countries (Foster and Rosenzweig, 1995; Conley and Udry, 2010), there is also growing evidence that manufacturing firms in developing countries learn from other firms. The learning spillovers may occur through suppliers, buyers, peers, or workers, among other channels.

Learning from suppliers was discussed briefly above in the context of the FDI spillovers literature. There is also evidence of learning through suppliers shared with foreign firms. Using a survey of Bangladeshi garment firms that elicited the top three suppliers of each firm, Kee (2015) finds that local “siblings” of foreign-owned firms, which share a local supplier, increased productivity and product scope when for arguably exogenous reasons the market share of the foreign-owned sibling expanded. Although these effects could simply reflect greater availability of particular types of
inputs, Kee suggests that the most important channel is knowledge flows. As noted above, Fieler et al. (2018) argue that quality upgrading by some producers can lead to quality upgrading by nearby firms that share suppliers.

Studies on learning from selling to foreign buyers or to locally based multinationals were discussed in Section 3.1 above. To date, it appears that there have been few studies in developing countries of learning from buyers who are not multinationals or on the export market. Evaluating the magnitude of spillovers from domestic buyers versus international buyers, and how these relate to product quality, seems to be a promising area for research.

Learning spillovers from peers, widely believed to exist, are challenging to document empirically, in part because of thorny econometric problems in estimating social effects (Manski, 1993). But recent studies have been able to manipulate experimentally the peer groups of entrepreneurs, to gain leverage for econometric identification. In an important contribution, Cai and Szeidl (2017) randomly assigned managers from 2,820 Chinese firms into groups that met monthly for one year. The meetings had a large effect on firm revenues (8.1%) and also had positive effects on profits and a management practice index similar to the World Management Survey score. To explore the learning channel directly, the authors randomly allocated information about a government grant and a high-return savings opportunity for managers, and found that not-directly-informed managers in groups where others had received the information were more likely to apply for both programs than not-directly-informed managers in groups where others had not received the information. In addition, they find that information about the government grant, which was plausibly perceived as more rival than the savings opportunity, was less likely to spill over when more firms in the group were direct competitors. No such difference is evident for the manager savings opportunity, which was less rival. Together, the results provide compelling evidence of learning spillovers between firms.

The Cai and Szeidl (2017) results contrast somewhat with a similar, earlier intervention by Fafchamps and Quinn (2018). By randomly assigning local entrepreneurs as judges in business-plan competitions in Ethiopia, Tanzania, and Zambia, Fafchamps and Quinn successfully generated in experimental variation in the judges’ peer networks. But the effects overall were quite modest. The authors found no effects on diffusion of management practices, client and supplier relations, or innovation, although they did find effects on tax registration and having a bank account (correcting for multiple hypothesis testing). The contrast with the Cai and Szeidl (2017) study is likely due in part to differences in the intensity of the peer interactions (in Fafchamps and Quinn (2018), the entrepreneurs met only once, rather than monthly for a year as in Cai and Szeidl (2017)) and in part to sample size (345 entrepreneurs in Fafchamps and Quinn (2018), 2,820 in Cai and Szeidl (2017)).

Two other notable recent studies have explored learning from peer firms in an experimental or quasi-experimental setting. Hardy and McCasland (2016) randomly allocated a new technology for weaving garments and training in using the technology and and they experimentally generated demand for products that required the technology. As in Cai and Szeidl (2017), they find that entrepreneurs are more likely to share information when they face less head-to-head competition. Although not focused on developing-country firms, Giorcelli (2019) is one of the few studies able to examine long-term outcomes of exposure to other firms. Under the Marshall plan in the 1950s, the US government sponsored trips of Italian managers to US firms and subsidized purchases by the Italian firms of advanced US technology. Giorcelli compares the set of firms that participated in the program to a set of firms that applied and were accepted but because of subsequent budget cuts were not able to participate. The sales, employment, and productivity of firms that participated
in the trips rose quickly and continued to rise steadily for at least 15 years. The productivity of firms that only received the technology subsidies also rose but reached a plateau after ten years. Outcomes for firms that received both were significantly greater than the sum of the effects for each alone, suggesting that there were complementarities between the trips and the technology subsidies.

Another channel through which firms may learn from other firms is employee flows. In one famous example, employees of a single Bangladeshi garment firm, Desh Garment Company, a joint venture with Daewoo Corporation, were sent to Korea for training in production techniques. More than 100 Korea-trained Desh employees subsequently moved to new firms, in many cases starting their own firms. These flows were an important catalyst for the growth of the Bangladeshi garment sector (Rhee, 1990; Rhee and Belot, 1990; Mostafa and Klepper, 2018). Recent papers have provided evidence on several types of spillovers through worker flows, although not (for the most part) on upgrading outcomes. Using Brazilian employer-employee data, Poole (2013) finds that when Brazilian firms hire workers who have previously worked in an MNC, the wages of incumbent workers rise.\(^{62}\) Researchers have also found evidence that employee movements lead “receiving” firms to export to similar destinations (e.g. Mion and Opremolla (2014) and Mion et al. (2016) in Portugal) and import from similar origins (e.g. Bisztray et al. (2018) in Hungary) as “sending” firms. Econometric identification of spillovers is always a challenge, but the accumulation of consistent findings raises one’s confidence that worker flows are an important channel for knowledge flows.

### 3.3.4.3 Learning from trainers/consultants

In addition to learning from their own experiences and learning from other enterprises, firms can also learn from trainers and consultants, whether their services are subsidized by governments or NGOs or purchased at market prices. An influential review of training experiments by McKenzie and Woodruff (2013), focused on small and medium-sized enterprises (SMEs), finds that most studies have very wide confidence intervals, with the result that it is rarely possible to reject a null hypothesis of no impact.\(^{63}\) (See also the reviews by Grimm and Paffhausen (2015) and Quinn and Woodruff (forthcoming). Because the literature has been thoroughly discussed in these previous reviews, here I will primarily focus on a few contributions that seem particularly relevant.

Bruhn et al. (2018) randomly allocated heavily subsidized consulting services, provided by private consulting firms, to SMEs (average employment: 14) in Puebla, Mexico. The intervention was of moderately high intensity: the firms met one-on-one with consultants for four hours per week for one year. There was not a uniform body of advice given; the consultants tailored their messages to the needs of the individual firms. The authors estimate positive short-term effects on productivity and return on assets, although these effects are only marginally significant (at the 10% level) and not robust in all specifications. By linking the experimental sample to administrative data from the Mexican social security agency, the authors were able to document significant effects on employment over a longer term (5 years).

Perhaps the most influential contribution in this area has been the consulting experiment of Bloom et al. (2013) in 17 Indian textile firms. The intervention was intensive: it provided one month of consulting from a multinational consulting firm to both treatment and control firms (the “diagnostic phase”) and then four months of consulting to treatment firms only (the “implementation-
tion phase"). The market value of the consulting services for the treated plants was approximately $250,000 USD per firm. The authors tracked 38 specific management practices, including performing regular maintenance on machines, tracking inventories at least weekly, monitoring quality defects daily, and offering performance pay to non-managerial and managerial staff. Using several methods to address concerns about small sample size, the authors find clear evidence that the implementation-phase consulting was effective both in increasing the share of the 38 management practices that firms adopted and in improving firm performance, measured in terms of output, TFP, or reductions of quality defects and inventory. The authors also use the consulting treatment as an instrument for the share of the 38 management practices adopted, to estimate the effect of the practices on performance (output, TFP, quality, inventory) and find significant coefficients on the management-practices variable. In a follow-up paper, Bloom et al. (forthcoming) find that the effects were still present nine years later: firms treated in the original experiment continued to employ more of the management practices, had greater worker productivity and higher-quality looms, and were more likely to be exporters.

This project has broken significant new ground in the study of firm behavior, and has rightfully been influential. But three notes of caution are in order. First, to interpret the instrumental-variables (IV) results as evidence for a causal effect of the specified management practices requires the exclusion restriction that the consulting affected performance only through its effect on the share of the 38 management practices adopted. If one believes that the four months of intensive consulting had effects on firm behavior that are not captured by the share-of-the-38-practices variable, then one should not interpret the IV estimates as causal effects of the management practices themselves. For this reason, this study should arguably not be considered definitive evidence for the “vertical” view, discussed in Section 2.2.3 above, that the 38 practices (or some subset of them) are better than existing practices across contexts. Note that this exclusion-restriction concern does not apply to the reduced-form (Intent-to-Treat) estimates of the effects of the consulting itself on performance, which are compelling. Second, the returns to the intervention are imprecisely measured. The authors did not have access to internal accounting data from the firms, and instead estimated profits based on their own performance estimates and a series of assumptions about the cost savings from reduction of waste fabric, profits expected to be derived from increased output, and other factors. On the basis of these assumptions, they estimate a return of $325,000 USD per year on the $250,000 USD worth of consulting services. Estimating profits in this way is an inexact science, and there is likely to be both significant heterogeneity and significant ex ante uncertainty in the profit effects. Third, relatedly, it is not clear that firms were making mistakes by not adopting the management practices on their own. Although the authors themselves are careful to attribute the lack of adoption to a lack of information, the paper appears to have been interpreted by others as showing that firms left money on the table, since the management practices themselves were cheap to implement (about $3,000 USD). But if we interpret the cost of consulting as part of the cost of adopting the new management practices, and allow for heterogeneity and uncertainty in the returns, then it is not obvious that firms left money on the table.

64The Atkin et al. (2017b) soccer-ball study provides one example where performance pay (in the form of piece rates) got in the way of technology adoption, and a less high-powered incentive scheme appeared to be more conducive to learning. See also Verhoogen (2016).

65The follow-up paper, Bloom et al. (forthcoming), was unfortunately unable to measure profits or productivity.

66In the notation of Section 2.1, the costs of acquiring knowledge and capabilities (I^k, I^K, and I^λ) may be sufficiently large that it is not worthwhile for the firm to incur them, given the heterogeneous and uncertain benefits. Recent work by Alfaro-Serrano (2019) emphasizes these costs of adoption and shows that a Peruvian program to
The Bloom et al. (2013) intervention was expensive, and it is worth investigating whether similar outcomes can be achieved more cheaply. Partnering with the Colombian government and focusing on autoparts firms, Iacoone et al. (2019) do this by comparing an intervention involving one-on-one consulting provided by local consultants (as opposed to more-expensive international consultants) to an intervention involving group consulting. The aim of the group consulting was to reduce costs and to take advantage of firms learning from one another. The authors find that both interventions had an effect on management practices, and that the group-consulting intervention (but not the individual consulting) had positive effects on employment and sales. Neither intervention had a significant positive effect on productivity, although the confidence bands are wide. Given that the group-consulting intervention is less costly, the study suggests that it would be the preferable design for scaling up.

The literature on training and consulting interventions is growing quickly. Several notable recent papers find positive effects on firm performance. Higuchi et al. (2017) randomized classroom and on-site training to 312 small manufacturers in Vietnam (average employment: 20), tracking firms over five years, and find positive effects on survival, sales, value-added, and profit. Higuchi et al. (2019) randomized classroom and on-site training, including quality control and production management practices as well as more standard topics such as marketing and record-keeping, to 113 small garment manufacturers in Tanzania (average employment: 5) and find positive effects on sales, value-added, and the number of products sold after 3 years. (See also Higuchi et al. (forthcoming).) Anderson et al. (2018) randomized marketing and financial skills training across 852 small enterprises in South Africa (average employment: 2.4), and find positive effects on profits, sales, and employment among the marketing group and on profits and cost-reductions among the financial-skills group.

Overall, although several studies have documented positive impacts, the effects of training and consulting interventions appear to be sensitive to the content of the advice and the details of implementation. The most successful interventions have tailored advice to the particular needs of firms, rather than providing cookie-cutter guidelines. It has often been important to follow firms over several years to see significant effects. The most successful interventions have been intensive, and in several cases expensive. Questions remain about whether firms leave money on the table by not purchasing training or consulting services and about which approaches are most cost-effective. At the same time, it seems clear that training and consulting can have significant positive effects on firm performance.

4 Conclusion

This paper has reviewed recent firm-level evidence on the drivers of upgrading in manufacturing firms in developing countries. From a measurement perspective, the literature faces a number of challenges. TFP measures have the conceptual advantage that they aim directly at capturing firm capabilities, but they have a number of well-known shortcomings. I have argued that focusing on directly observable measures of upgrading — technology use, quality ratings, expansions of product scope, and productivity measured under controlled conditions — is a particularly promising way forward. At the same time, such measures are often available only for particular sectors, and increases in these measures are not necessarily optimal either for firms or for the broader economies subsidize certifications such as ISO 9001, which require formalization and documentation of processes but not particular management practices, had the indirect effect of increasing adoption of higher-scoring management practices.
in which they are embedded. It seems clear that there is value in improving measurement both of indirect measures such as TFP and of more direct measures of upgrading.

Despite the difficulties of measurement, several empirical patterns emerge. Selling to developed-country consumers, either directly or indirectly through value chains with richer-country end-consumers, appears to be robustly associated with upgrading. Increased availability of high-quality inputs also appears to promote upgrading. It is not clear that developing-country firms are making mistakes by not upgrading but there is growing evidence that tailored, intensive consulting interventions can improve firm performance. A broader conclusion is that developing-country firms appear to be constrained by a lack of know-how — both of internal capabilities and knowledge of products and techniques. A key challenge, perhaps the key challenge, in promoting upgrading is to promote learning by firms.

A number of important questions remain open. One is the strength of the link between the products that a firm specializes in and the rate of learning. Does producing higher-quality products, for instance, generate a greater accumulation of know-how? The hypothesis that there is a link between the pattern of specialization and upgrading was central to the thinking of an early generation of development economists (e.g. Prebisch (1950)). In recent years, it has been advanced by Dani Rodrik, Ricardo Hausman and others (see e.g. Hausmann et al. (2007), Hausmann et al. (2014)) and investigated largely at the sectoral level. Now that firm-product-level datasets are increasingly available, the time seems ripe for investigating the link at the firm level.

Another important question is to what extent behavioral biases of entrepreneurs lead them not to maximize profits. I have argued that what appears to be non-optimizing behavior by firms can often be explained by firms’ lack of know-how, organizational dynamics, or other constraints, and that we need to think carefully about these possibilities before concluding that individuals are failing to optimize. But firm owners and managers are human, and they may procrastinate, put more weight on losses than gains, ignore evidence that does not comport with their priors, and display all the other foibles that other humans do. There is a need for research designs that can measure these propensities separately from other factors.

Also important is to what extent firm capabilities can be acquired on markets or must be homegrown. In principle, one would expect firms to be able to hire consultants to acquire the know-how needed to upgrade. One puzzling fact, worthy of further investigation, is that in many developing-country settings the consulting market is either extremely thin or non-existent. But even where consulting markets exist, it requires time and effort for firms to incorporate new knowledge or practices into the everyday functioning of an organization. A related question is to what extent firms can improve their performance by hiring highly skilled managers, even if their homegrown capabilities are low. In many developing countries, the supply of highly skilled managers is limited. But it also appears that top managers cannot just parachute in and impose new practices; the capabilities to implement practices effectively must be developed internally as well. More research on these issues is much needed.

Finally, it is natural to ask about the implications of the recent literature on upgrading for industrial policy. This review has focused on the determinants of upgrading behavior by firms, with the idea that such an understanding will eventually be useful in policy design. But policymakers must face a number of additional constraints not considered here, among them pressures from different interest groups and the limited knowledge of government officials. More research is needed on what works and what does not work in industrial policy, especially given the limited capacity of many developing-country governments. If policies are to be implemented at scale, designers will also need to confront the general-equilibrium effects of industrial-policy interventions, which have
not been the focus here. Analyzing these issues will likely require more guidance from economic
theory than the primarily reduced-form studies discussed in this review have relied on.

Although much work obviously remains to be done, there are many reasons for optimism
about the prospects for future research on firm-level upgrading in developing countries. The
data frontier has been expanding quickly, with information on customs transactions, firm-to-firm
trade, quantities and prices at the product level, banking relationships, and other sorts of contracts
becoming increasingly available. Appreciation is growing in a number of fields — macroeconomics,
industrial organization, and international trade, as well as development — for careful firm-level
empirical work on the determinants of innovative behavior. And policymakers in many countries
are hungry for rigorous, evidence-based advice about how to promote upgrading. It is an exciting
time for the field.
References


, and , “Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending


Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings,” *Journal of Political Econ-


Jack, B. Kelsey, “Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries,” 2013. Agricultural Technology Adoption Initiative, J-PAL (MIT) and CEGA (Berkeley).


Labanca, Claudio, Danielen Molina, and Marc-Andreas Muendler, “Preparing to Export,” 2014. Unpub. paper, UCSD.


Levitt, Steven D., John A. List, and Chad Syverson, “Toward an Understanding of Learning by
_ and _, “Improving Export Quality in the Colombian Coffee Chain,” 2019. Unpub. paper, LSE.


Poole, Jennifer P., “Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility,” *Review of Economics and Statistics*, 2013, 95 (2), 393-406.


Rhee, Yung Whee, “The Catalyst Model of Development: Lessons from Bangladesh’s Success with


forthcoming.

Teshima, Kensuke, “Import Competition and Innovation at the Plant Level: Evidence from Mexico,” 2010. Unpub. paper, ITAM.


Tybout, James, “Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?,” *Journal of Economic Literature*, 2000, 38, 11–44.


