What Determines Productivity?*

Chad Syverson
University of Chicago Booth School of Business
and National Bureau of Economic Research
chad.syverson@chicagobooth.edu

April 2010

Abstract

Economists have shown that large and persistent differences in productivity levels across businesses are ubiquitous. This finding has shaped research agendas in a number of fields, including (but not limited to) macroeconomics, industrial organization, labor, and trade. This paper surveys and evaluates recent empirical work addressing the question of *why* businesses differ in their measured productivity levels. The causes are manifold, and differ depending on the particular setting. They include elements sourced in production practices—and therefore over which producers have some direct control, at least in theory—as well as from producers' external operating environments. After evaluating the current state of knowledge, I lay out what I see are the major questions that research in the area should address going forward.

* I thank Eric Bartelsman, Nick Bloom, Roger Gordon, John Haltiwanger, Chang-Tai Hsieh, Ariel Pakes, Amil Petrin, John Van Reenen, and anonymous referees for helpful comments. This work is supported by the NSF (SES-0519062 and SES-0820307), and both the Stigler Center and the Centel Foundation/Robert P. Reuss Faculty Research Fund at the University of Chicago Booth School of Business. Contact information: University of Chicago Booth School of Business; 5807 S. Woodlawn Ave.; Chicago, IL 60637.

1. Introduction

Thanks to the massive infusion of detailed production activity data into economic study over the past couple of decades, researchers in many fields have learned a great deal about how firms turn inputs into outputs. Productivity, the efficiency with which this conversion occurs, has been a topic of particular interest. The particulars of these studies have varied depending on the researchers' specific interests, but there is a common thread. They have documented, virtually without exception, enormous and persistent measured productivity differences across producers, even within narrowly defined industries.

The magnitudes involved are striking. I found in Syverson (2004a) that within 4-digit SIC industries in the U.S. manufacturing sector, the average difference in logged total factor productivity (TFP) between an industry's 90^{th} and 10^{th} percentile plants is 0.651. This corresponds to a TFP ratio of $e^{0.651} = 1.92$. To emphasize just what this number implies, it says that the plant at the 90^{th} percentile of the productivity distribution makes almost *twice* as much output with the *same measured inputs* as the 10^{th} percentile plant. Note that this is the average 90-10 range. The range's standard deviation across 4-digit industries is 0.173, so several industries see much larger productivity differences among their producers. U.S. manufacturing is not exceptional in terms of productivity dispersion. Indeed, if anything, it is small relative to the productivity variation observed elsewhere. Chang-Tai Hsieh and Peter J. Klenow (2009), for example, find even larger productivity differences in China and India, with average 90-10 TFP ratios over 5:1.[2]

These productivity differences across producers are not fleeting, either. Regressing a producer's current TFP on its one-year-lagged TFP yields autoregressive coefficients on the order of 0.6 to 0.8 (see, e.g., Árpád Ábrahám and Kirk White (2006) and Foster, Haltiwanger, and Syverson (2008)). Put simply, some producers seem to have figured out their business (or at least are on their way), while others are woefully lacking. Far more than bragging rights are at

[2] These figures are for revenue-based productivity measures; i.e., where output is measured using plant revenues (deflated across years using industry-specific price indexes). TFP measures that use physical quantities as output measures rather than revenues actually exhibit even *more* variation than do revenue-based measures, as documented in Lucia Foster, John Haltiwanger, and Syverson (2008). Hsieh and Klenow (2009) also find greater productivity dispersion in their TFP measures that use quantity proxies to measure output (actual physical quantities are not available for most producers in their data). Even though it is only a component of revenue-based TFP (the other being the producer's average price), quantity-based TFP can be more dispersed because it tends to be negatively correlated with prices, as more efficient producers sell at lower prices. Thus revenue-based productivity measures, which combine quantity-based productivity and prices, tend to understate the variation in producers' physical efficiencies.

stake here: another robust finding in the literature—virtually invariant to country, time period, or industry—is that higher productivity producers are more likely to survive than their less efficient industry competitors. Productivity is quite literally a matter of survival for businesses.

1.1. How Micro-Level Productivity Variation and Persistence Has Influenced Research

The discovery of ubiquitous, large, and persistent productivity differences has shaped research agendas in a number of fields. Here are some examples of this influence, though by no means is it meant to be a comprehensive accounting. They speak to the breadth of the impact that answers to this paper's title question would have.

Macroeconomists are dissecting aggregate productivity growth—the source of almost all per capita income differences across countries—into various micro-components, with the intent of better understanding the sources of such growth. Foster, Halitwanger, and C.J. Krizan (2001), for example, overview the substantial role of reallocations of economic activity toward higher productivity producers (both among existing plants and through entry and exit) in explaining aggregate productivity growth. Hsieh and Klenow (2009) ask how much larger the Chinese and Indian economies would be if they achieved the same efficiency in allocating inputs across production units as does the U.S. Models of economic fluctuations driven by productivity shocks are increasingly being enriched to account for micro-level patterns, and are estimated and tested using plant- or firm-level productivity data rather than aggregates (e.g., Jeffrey R. Campbell and Jonas D. M. Fisher (2004); Eric Bartelsman, Haltiwanger, and Stefano Scarpetta (2008); Marcelo Veracierto (2008)). Micro productivity data have also been brought to bear on issues of long-run growth, income convergence, and technology spillovers. They offer a level of resolution unattainable with aggregated data.

In industrial organization, research has linked productivity levels to a number of features of technology, demand, and market structure. Examples include the effect of competition (Syverson (2004b), James A. Schmitz (2005)), the size of sunk costs (Allan Collard-Wexler (2008)), and the interaction of product market rivalry and technology spillovers (Nicholas Bloom, Mark Schankerman, and John Van Reenen (2007)). Another line of study has looked at the interaction of firms' organizational structures with productivity levels (e.g., Vojislav Maksimovic and Gordon Phillips (2002), and Antoinette Schoar (2002), and Ali Hortaçsu and Syverson (2007, 2009)).

Labor economists have explored the importance of workers' human capital in explaining productivity differences (John Abowd, Haltiwanger, Ron Jarmin, Julia Lane, Paul Lengermann, Kristin McCue, Kevin McKinney, and Kristin Sandusky (2005) and Jeremy Fox and Valerie Smeets (2009)), the productivity effects of incentive pay (Edward P. Lazear (2000)), other various human resources practices (Casey Ichniowski and Kathryn Shaw (2003)), managerial talent and practices (Bloom and Van Reenen (2007)), organizational form (Luis Garicano and Paul Heaton (2007)), and social connections among co-workers (Oriana Bandiera, Iwan Barankay, and Imran Rasul (2009)). There has also been a focus on the role of productivity-driven reallocation on labor market dynamics via job creation and destruction (Haltiwanger, Scarpetta, and Helena Schweiger (2008)).

Perhaps in no other field have the productivity dispersion patterns noted above had a greater influence on the trajectory of the research agenda than in the trade literature. Theoretical frameworks using heterogeneous-productivity firms like Jonathan Eaton and Samuel Kortum (2002) and Marc J. Melitz (2003) are now the dominant conceptual lenses through which economists view trade impacts. In these models, the trade impacts vary across producers, and depend on their productivity levels in particular. Aggregate productivity gains come from improved selection and heighted competition that trade brings. A multitude of empirical studies have accompanied and been spurred by these theories (e.g., Nina Pavcnik (2002); Andrew Bernard, J. Bradford Jensen, and Peter Schott (2006); and Eric A. Verhoogen (2008)). They have confirmed many of the predicted patterns, and raised questions of their own.

1.2. The Question of "Why?"

Given the important role that productivity differences play in these disparate literatures, the facts above raise obvious and crucial questions. *Why* do firms (or factories, stores, offices, or even individual production lines, for that matter) differ so much in their abilities to convert inputs into output? Is it dumb luck, or instead something—or many things—more systematic? Can producers control the factors that influence productivity, or are they purely external products of the operating environment? What supports such large productivity differences in equilibrium?

A decade ago, when Bartelsman and Mark Doms (2000) penned the first survey of the micro-data productivity literature for this journal, researchers were just beginning to ask the "Why?" question. Much of the work to that point had focused on establishing facts like those above—the "What?" of productivity dispersion. Since then, the literature has focused more intensely on the reasons why productivity levels are so different across businesses. There has definitely been progress. But we've also learned more about what we *don't* know, and this is guiding the ways in which the productivity literature will be moving. This article is meant to be a guide to and comment on this research.

I begin by setting some boundaries. I have to. A comprehensive overview of micro-founded productivity research is neither possible in this format nor desirable. There are simply too many studies to allow adequate coverage of each. First, I will focus on empirical work. This is not because I view it as more important than theory. Rather, it affords a deeper coverage of this important facet of a giant literature, and it better reflects my expertise. That said, I will sketch out a simple heterogeneous-productivity industry model below to focus the discussion, and I will also occasionally bring up specific theoretical work with particularly close ties to the empirical issues discussed. Furthermore, for obvious reasons, I will focus on research that has been done since Bartelsman and Doms (2000) was written.

Even within these boundaries, there are more studies than can be satisfactorily described individually. I see this article's role as filtering the broader lessons of the literature through the lens of a subset of key studies. The papers I focus on here are not necessarily chosen because they are the first or only good work on their subject matter, but rather because they had an archetypal quality that lets me weave a narrative of the literature. I urge readers whose interests have been piqued to more intensively explore the relevant literatures. There is far more to be learned than I can convey here.

A disclaimer: some of my discussion contains elements of commentary. These opinions are mine alone and may not be the consensus of researchers in the field.

I organize this article as follows. The next section sketches the conceptual background: what productivity is, how it is often measured in practice, and how differences in productivity among producers of similar goods might be sustained in equilibrium. Section 3 looks at influences on productivity that operate primarily within the business. This can be at the firm level, plant level, or even on specific processes within the firm. Many of these influences may potentially be under the control of the economic actors inside the business. In other words, they can be "levers" that management or others have available to impact productivity. Section 4 focuses on the interaction of producers' productivity levels and the markets in which they operate. These are elements of businesses' external environments that can affect productivity levels. This impact might not always be direct, but they can induce producers to pull some of the levers discussed in Section 3, indirectly influencing observed productivity levels in the process. They may also be factors that affect the amount of productivity dispersion that can be sustained in equilibrium, and influence observed productivity differences through that channel. Section 5 discusses what I see as the big questions about business-level productivity patterns that still need to be answered. A short concluding section follows.

2. Productivity—What It Is, How It Is Measured, and How Its Dispersion Is Sustained

This section briefly reviews what productivity is conceptually, how it is measured in practice, and how productivity differences among producers of similar goods might be supported in equilibrium. Deeper discussions on the theory of productivity indexes can be found in Douglas W. Caves, Laurits R. Christensen, and W. Erwin Diewert (1982) and the references therein. More detail on measurement issues can be found in the large literature on the subject; see, for example, G. Steven Olley and Ariel Pakes (1996); Zvi Griliches and Jacques Mairesse (1998); Richard Blundell and Stephen R. Bond (2000); James Levinsohn and Amil Petrin (2003); and Daniel Ackerberg, C. Lanier Benkard, Steven Berry, and Ariel Pakes (2007). Examples of models that derive industry equilibria with heterogeneous-productivity producers include Boyan Jovanovic (1982); Hugo A. Hopenhayn (1992); Richard Ericson and Pakes (1995); Melitz (2003); Marcus Asplund and Volker Nocke (2006); and Foster, Haltiwanger, and Syverson (2008).

2.1. Productivity in Concept

Simply put, productivity is efficiency in production: how much output is obtained from a given set of inputs. As such, it is typically expressed as an output-input ratio. Single-factor productivity measures reflect units of output produced per unit of a particular input. Labor productivity is the most common measure of this type, though occasionally capital or even materials productivity measures are used. Of course, single-factor productivity levels are affected by the intensity of use of the excluded inputs. Two producers may have quite different labor productivity levels even though they have the same production technology, if one happens to use capital much more intensively, say because they face different factor prices.

Because of this, researchers often use a productivity concept that is invariant to the intensity of use of observable factor inputs. This measure is called total factor productivity, or TFP (it is also sometimes called multifactor productivity, MFP). Conceptually, TFP differences reflect shifts in the isoquants of a production function: variation in output produced from a fixed set of inputs. Higher-TFP producers will produce greater amounts of output with the same set of observable inputs than lower-TFP businesses, and hence have isoquants that are shifted down and to the left. Factor price variation that drives factor intensity differences does not affect TFP, because it induces shifts *along* isoquants rather than shifts *in* isoquants.

TFP is most easily seen in the often-used formulation of a production function where output is the product of a function of observable inputs and a factor-neutral (alternatively, Hicks-neutral) shifter:

$$Y_t = A_t F(K_t, L_t, M_t),$$

where Y_t is output, $F(\cdot)$ is a function of observable inputs capital Y_t, labor L_t, and intermediate materials M_t, and A_t is the factor-neutral shifter. In this type of formulation, TFP is A_t. It captures variations in output not explained by shifts in the observable inputs that act through $F(\cdot)$.[3]

TFP is, at its heart, a residual. As with all residuals, it is in some ways a measure of our ignorance: it is the variation in output that cannot be explained based on observable inputs. So it is fair to interpret the work discussed in this survey as an attempt to "put a face on" that residual—or more accurately, "put faces on," given the multiple sources of productivity variation. The literature has made progress when it can explain systematic influences on output across production units that do not come from changes in observable inputs like standard labor or capital measures.

2.2. Measuring Productivity

While productivity is relatively straightforward in concept, a host of measurement issues arise when constructing productivity measures from actual production data. Ironically, while

---

[3] I use a multiplicatively separable technology shift to make exposition easy, but TFP can be extracted from a general time-varying production function $Y_t = G_t(A_t,K_t,L_t,M_t,)$. Totally differentiating this production function gives:

$$dY_t = \frac{\partial G}{\partial A} dA_t + \frac{\partial G}{\partial K} dK_t + \frac{\partial G}{\partial L} dL_t + \frac{\partial G}{\partial M} dM_t.$$

Without loss of generality, we can choose units to normalize $\partial G/\partial A = 1$. Thus when observed inputs are fixed ($dK_t = dL_t = dM_t = 0$), differential shifts in TFP, dA_t, create changes in output dY_t.

research with micro production data greatly expands the set of answerable question and moves the level of analysis closer to where economic decisions are made than aggregate data does, it also raises measurement and data quality issues more frequently.

The first set of issues regards the output measure. Many businesses produce more than one output. Should these be aggregated to a single output measure, and how if so? Further, even detailed producer microdata do not typically contain measures of output quantities. Revenues are typically observed instead. Given this limitation of the data, the standard approach has been to use revenues (deflated to a common year's real values using price deflator series) to measure output. While this may be acceptable, and even desirable, if product quality differences are fully reflected in prices, it can be problematic whenever price variation instead embodies differences in market power across producers. In that case, producers' measured productivity levels may reflect less about how efficient they are and more about the state of their local output market. Recent work has begun to dig deeper into the consequences of assuming single-product producers and using revenue to measure output. I'll discuss this more below. In the mean time, I will go forward assuming deflated revenues accurately reflect the producer's output.

The second set of measurement issues regards inputs. For labor, there is the choice of whether to use number of employees, employee-hours, or some quality-adjusted labor measure (the wage bill is often used in this last role, based on the notion that wages capture marginal products of heterogeneous labor units). Capital is typically measured using the establishment or firm's book value of its capital stock. This raises several questions. How good of a proxy is capital stock for the flow of capital services? Should the stock be simply the producer's reported book value, and what are the deflators? Or should the stock be constructed using observed investments and the perpetual inventory method—and what to assume about depreciation? When measuring intermediate materials, an issue similar to the revenue-as-output matter above arises, because typically only the producer's total expenditures on inputs are available, not input quantities. More fundamentally, how should intermediate inputs be handled? Should one use a gross output production function and include intermediate inputs directly, or should intermediates simply be subtracted from output so as to deal with a value-added production function? On top of all these considerations, one makes these input measurement choices in the context of knowing that any output driven by unmeasured input variations (due to input quality differences or intangible capital, for example) will show up as productivity.

The third set of measurement concerns involves aggregating multiple inputs in a TFP measure. As described above, TFP differences reflect shifts in output while holding inputs constant. To construct the output-input ratio that measures TFP, a researcher must weight the individual inputs appropriately when constructing a single-dimensional input index. The correct weighting is easiest to see when the production function is Cobb-Douglas:

$$TFP_t = A_t = \frac{Y_t}{K_t^{\alpha_k} L_t^{\alpha_l} M_t^{\alpha_m}}.$$

In this case, the inputs are aggregated by taking the exponent of each factor to its respective output elasticity. It turns out that this holds more generally as a first-order approximation to any production function. The input index in the TFP denominator can be constructed similarly for general production functions.[4]

Even after determining how to construct the input index, one must measure the output elasticities α_j, $j \in \{k,l,m\}$. Several approaches are common in the literature. One builds upon assumptions of cost-minimization to construct the elasticities directly from observed production data. A cost-minimizing producer will equate an input's output elasticity with the product of that input's cost share and the scale elasticity. If cost shares can be measured (obtaining capital costs are usually the practical sticking point here), and the scale elasticity either estimated or assumed, then the output elasticities α_j can be directly constructed. If a researcher is willing to make some additional but not innocuous assumptions—namely, perfect competition and constant returns to scale—then the elasticities equal the share of revenues paid to each input. This makes constructing the α_j simple. Materials' and labor's shares are typically straightforward to collect with the wage bill and materials expenditures data at hand. Capital's share can be constructed as the residual, obviating the need for capital cost measures. (Though there is a conceptual problem since, as the model that follows below points out, it is unclear what makes the producer's size finite in a perfectly competitive, constant returns world.) An important caveat is that the index approach assumes away factor adjustment costs. If they are present, the first-order conditions linking observed factor shares to output elasticities will not hold. This can be mitigated in part (but at cost) by using cost shares that have been averaged over either time or producers in order

[4] While Cobb-Douglas-style approaches are probably the most common in the literature, many researchers also use the translog form (see Caves, Christensen, and Diewert 1982), which is a second-order approximation to general production functions, and as such is more flexible, though more demanding of the data. There is also an entirely nonparametric approach, data envelopment analysis (DEA), that is used in certain, somewhat distinct circles of the literature. See William W. Cooper, Lawrence M. Seiford, and Kaoru Tone (2006) for an overview of DEA methods.

to smooth out idiosyncratic adjustment-cost-driven misalignments in actual and optimal input levels, but some mismeasurement could remain.

A separate approach is to estimate the elasticities α_j by estimating the production function. In this case, (logged) TFP is simply the estimated sum of the constant and the residual. In the Cobb-Douglas case (which again, recall, is a first-order approximation to more general technologies), the estimated equation is:

$$\ln Y_t = \alpha_0 + \alpha_k \ln K_t + \alpha_l \ln L_t + \alpha_m \ln M_t + \omega_t.$$

Hence the TFP estimate would be $\hat{\alpha}_0 + \widehat{\omega}_t$, where the first term is common across production units in the sample (typically the technology is estimated at the industry level), and the second is idiosyncratic to a particular producer.

This approach raises econometric issues. As first pointed out by Marschak and Andrews (1944), input choices are likely to be correlated with the producer's productivity ω_t: more efficient producers are, all else equal, likely to hire more inputs. There is also potential selection bias when a panel is used, since less efficient producers—those with low ω_t—are more likely to exit from the sample. (As will be discussed below, the positive correlation between productivity and survival is one of the most robust findings in the literature.) Then there is the issue of producer-level price variation mentioned above. A substantial literature has arisen to address these issues; see Griliches and Mairesse (1998); Ackerberg, Benkard, Berry, and Pakes (2007); and Johannes Van Biesebroeck (2008) for overviews.

There is debate as to which of the many available methods is best. In the end, as I see it, choosing a method is a matter of asking oneself which assumptions one is comfortable making. Certainly one cannot escape the fact that *some* assumptions must be made when estimating the production function.

Fortunately, despite these many concerns, many of the results described in this paper are likely to be quite robust to measurement peculiarities. When studies have tested robustness directly, they typically find little sensitivity to measurement choices. The inherent variation in establishment- or firm-level microdata is typically so large as to swamp any small measurement-induced differences in productivity metrics. Simply put, high-productivity producers will tend to look efficient regardless of the specific way that their productivity is measured. I usually use cost-share-based TFP index numbers as a first pass in my own work; they are easy to construct and offer the robustness of being a nonparametric first-order approximation to a general production function. That said, it is always wise to check one's results for robustness to specifics of the measurement approach.

2.3. A Model of Within-Industry Productivity Dispersion

Given the large differences in productivity within an industry that I discussed above, a natural question is to ask how they could be sustained in equilibrium. The ubiquity of this dispersion suggests there must be some real economic force at work, rather than it simply being an artifact of measurement or odd chance. Here, I sketch out a simple model that shows how that is possible. The model will also prove helpful in facilitating discussion throughout this survey.

Industry producers, indexed by i, earn profits given by

$$\pi_i = R(A_i, L_i, D) - wL_i - f.$$

$R(\cdot)$ is a general revenue function. A_i is the producer's productivity level, and L_i is its labor input. (I assume labor is the firm's only input for the sake of simplicity.) Productivity levels differ across producers. The specific form of $R(\cdot)$ depends on the structure of the output market and the production function. Revenues can also depend on an industry state D. This can be a vector or a scalar, and depending on the structure of output market competition, it may include industry-wide demand shocks, the number of industry producers, their productivity levels, and/or moments of the productivity distribution. Both the wage rate w and fixed cost f are common across, and taken as given by, all producers.

I assume $R(\cdot)$ is twice differentiable with $\partial R/\partial L > 0$, $\partial^2 R/\partial L^2 < 0$, $\partial R/\partial A > 0$, and $\partial^2 R/\partial A\partial L > 0$. If the industry is perfectly competitive, these conditions are satisfied given a production function that is similarly differentiable, concave in L, and where productivity and labor are complements. Further, under perfect competition, all information contained in D is reflected in the market price P that equates total demand and supply, which the producers of course take as given. In imperfectly competitive markets, the assumptions about $R(\cdot)$ place restrictions on the form of competitive interaction (be it monopolistically competitive or oligopolistic) and through this the shapes of the residual demand curves. The contents of D will also depend on the particulars of the competitive structure. For example, in a heterogeneous-cost Cournot oligopoly, D will contain the parameters of the industry demand curve and the productivity levels of the industry's producers, as these are sufficient to determine the Nash equilibrium outputs and therefore revenues of each producer *i*. Despite these restrictions, this setup is reasonably general.

The assumptions on the shape of $R(\cdot)$ imply that given the industry state D, each producer has a unique optimal employment level L_i^* that is increasing in its productivity level. Intuitively, the producer's optimal employment level (which I refer to from here forward as its size), which is set to equate marginal revenues and marginal costs, is pinned down by increasing marginal costs in perfectly competitive markets and a downward-sloping residual demand curve (and possibly increasing marginal costs as well) in imperfectly competitive markets.

Denote the producer's profits at its optimal size by

$$\pi(A_i, L_i^*, D) = R(A_i, L_i^*, D) - wL_i^* - f.$$

By the envelope theorem and the conditions on the revenue function, profits are increasing in the producer's productivity level A_i. This implies that there will be a critical productivity level $\underline{A}$ such that for $A_i < \underline{A}$, profits will be negative. $\underline{A}$ will depend in general on w, f, and the industry state D. Since D may itself depend on the distribution of productivity levels in the industry, we will need an additional condition to determine the industry equilibrium. This comes from an entry structure as follows.

A large pool of ex-ante identical potential entrants decides whether to enter the industry. They first choose whether to pay a sunk entry cost s in order to receive a productivity draw from a distribution with probability density function $g(A)$ over the interval $[A_l, A_u]$.[5] If a potential entrant chooses to receive a draw, it determines after observing it whether to begin production at its optimal size and earn the corresponding operating profits $\pi(A_i,L_i^*,D)$.

Only potential entrants with productivity draws high enough to make nonnegative operating profits will choose to produce in equilibrium. Hence the expected value of paying s is the expected value of $\pi(A,L^*,D)$ over $g(A)$, conditional on drawing $A_i \geq \underline{A}$. This expected value is obviously affected by the cutoff cost level $\underline{A}$. A free-entry condition pins down this value: $\underline{A}$ must set the net expected value of entry into the industry V^e equal to zero. Thus $\underline{A}$ satisfies

$$V^e = \int_{\underline{A}}^{A_u} \pi(A, L^*, D) g(A) dA - s = 0.$$

[5] These bounds are essentially arbitrary as long as they span $\underline{A}$ for any possible D. That is, a producer with productivity level A_l is not profitable (i.e., it cannot cover its fixed costs) in any possible industry state, and one with productivity A_u is always profitable.

This expression summarizes the industry equilibrium.[6] It combines the two conditions that all producers make nonnegative operating profits and that entry occurs until the expected value of taking a productivity draw is zero. By pinning down the equilibrium distribution of productivity levels in the industry through determining $\underline{A}$, it also determines the equilibrium industry state *D*. The particular values of $\underline{A}$ and *D* depend on the exogenous components of the model: $g(A)$, w, f, and s, and the functional form of $R(\cdot)$.

The equilibrium productivity distribution will be a truncation of the underlying productivity distribution $g(A)$. Specifically, the equilibrium distribution (denoted $\gamma(A)$) is:

$$\gamma(A) = \begin{cases} \frac{g(A)}{1-G(\underline{A})} & \text{if } A \geq \underline{A} \\ 0 & \text{otherwise} \end{cases}.$$

There are two notable features of this distribution. First, it is not trivially degenerate; the model supports productivity heterogeneity under general conditions. This is because high-productivity producers are limited in their ability to sell to the industry's entire market. This finite optimal producer size is a consequence of the concavity of the revenue function. In perfectly competitive markets, this concavity comes from increasing marginal costs. In industries with imperfectly competitive output markets, the concavity arises from downward-sloping demand curves (due to product differentiation from any source) and, possibly, from increasing marginal costs as well. In either case, one can interpret productivity *A* as a factor of production that differs in quantity or quality across producers. A higher level of *A* loosens the size constraint but does not eliminate it.

Second, the average productivity level in the industry will vary as the exogenous parameters change. Increases in the average productivity level across plants (coming from parameter changes that increase $\underline{A}$) will thus expectedly translate into higher aggregate industry productivity—the ratio of total industry output to total industry inputs.[7] Therefore what happens

[6] I've made two implicit assumptions in this equation. First, V^e is exactly zero only in industries with a large number of producers. I will assume there is a continuum of producers for the remainder of the discussion. This is consistent with an assumption of perfect competition or monopolistic competition in the product market, though obviously rules out strategic oligopolistic interactions. The model's logic applies to industries with a discrete number of firms, however. In that case, free entry condition will imply a number of producers *N* such that the expected value of entry with *N* firms is positive but is negative with *N*+1 firms. The other assumption is that the productivity distribution g(A) is continuous, but the model can be modified to accommodate discrete productivity distributions.

[7] For differentiated product industries, relating an industry's aggregate productivity level to the productivity levels of its component firms requires constructing a quantity index that adds up firms' disparate outputs. The proper index depends on how the product varieties enter final demanders' utility functions. Under standard aggregators, increases in the average firm-level productivity translate into increases in aggregate industry productivity (see, e.g.,

at the micro level feeds upwards into aggregates. This feature reflects a major thrust behind the research agenda of understanding micro productivity: it teaches us more about aggregate productivity movements.

Of course, this model is very simple and leaves out many features observed in empirical work on productivity. I will quickly discuss two such features.

As a two-stage entry and production model, the model abstracts from dynamics. It can therefore be interpreted as characterizing long-run industry equilibria. That said, versions of this model's type with more complex dynamics have been worked out by, among others, Hopenhayn (1992) and Asplund and Nocke (2006). Further, even this simple structure hints at how the dynamics of reallocation—a focus of some of the literature discussed below—might work. Suppose the industry is initially in equilibrium and then each producer is hit with a persistent, independent productivity shock. Those receiving favorable shocks will see an increase in their optimal size, while those hit by negative shocks will want to shrink. Indeed, some may be hit by shocks so adverse that they will no longer be profitable. And if we imagine there are still potential entrants who could pay the sunk cost to take a productivity draw, the environment after the productivity shocks may be favorable enough to induce some of them to want to do so. Any adjustment to a new, post-shock equilibrium will therefore require reallocation of inputs from their initial locations. Favorably shocked producers will grow, unfavorably shocked producers will shrink or exit, and new producers may enter the industry at a productivity level above exiters. These patterns of reallocation are robust features of the data.

A greater limitation of the model is that a producer's productivity is exogenous. The equilibrium productivity distribution is endogenized only through a selection effect: the determination of who produces in equilibrium via $\underline{A}$. While I discuss below that selection is an empirically important mechanism, it is abundantly clear that producers often take actions to try to raise their productivity level. In this case, the equilibrium sketched out above will not directly apply, though many of its basic elements will.

Despite the model's simplicity and limited scope, it can form a useful conceptual base upon which to build the discussion below.

Melitz (2003)). However, there are complications involved in empirically mapping back-and-forth between changes in micro-level productivity distributions within an industry and changes in aggregate industry productivity (see, e.g., Organization for Economic Cooperation and Development (2003); Petrin and Levinsohn (2008); Susanto Basu, Luigi Pascali, Fabio Schiantarelli, and Luis Serven (2009); and Charles R. Hulten (2009)).

3. Productivity and the Plant or Firm

This section discusses factors that directly impact productivity at the micro level by operating within the plant or firm. They are "levers" that management or others can potentially use to impact the productivity of their business. They are akin to forces that would allow firms in the model of the previous section to raise their A_i draw, though most likely at a cost. Section 4 below will focus on influences external to the firm: elements of the industry or market environment that can induce productivity changes or support productivity dispersion.

I have broken up the discussion of direct productivity impacts by category for the sake of exposition. However, it's good to keep in mind that some forces can overlap these categories, and multiple mechanisms can act in concert. I will point out many of these across-category links as the discussion goes along.

3.1. Managerial Practice/Talent

Researchers have long proposed that managers drive productivity differences.[8] Whether sourced in the talents of the managers themselves or the quality of their practices, this is an appealing argument. Managers are conductors of an input orchestra. They coordinate the application of labor, capital, and intermediate inputs. Just as a poor conductor can lead to a cacophony rather than a symphony, one might expect poor management to lead to discordant production operations.

Still, perhaps no potential driver of productivity differences has seen a higher ratio of speculation to actual empirical study. Data limitations have been the stumbling block. The proliferation of production microdata has afforded a great increase in detail, but such data rarely

[8] I mean *long* proposed: Francis A. Walker (1887) posits that managerial ability is the source of differences in surplus across businesses: "The excess of produce which we are contemplating comes from directing force to its proper object by the simplest and shortest ways; from saving all unnecessary waste of materials and machinery; from boldly incurring the expense—the often large expense—of improved processes and appliances, while closely scrutinizing outgo and practising a thousand petty economies in unessential matters; from meeting the demands of the market most aptly and instantly; and, lastly, from exercising a sound judgment as to the time of sale and the terms of payment. It is on account of the wide range among the employers of labor, in the matter of ability to meet these exacting conditions of business success, that we have the phenomenon in every community and in every trade, in whatever state of the market, of some employers realizing no profits at all, while others are making fair profits; others, again, large profits; others, still, colossal profits." It is impressive how Walker's description closely matches (albeit with the flowing prose typical of the time) the viewpoints of researchers over 120 years later. We finally are becoming able, with the growing availability of broad-based production microdata, to test such hypotheses on a comprehensive basis.

contains detailed information on any aspect of managerial inputs. Sometimes there may be a distinction made between blue- and white-collar or production and nonproduction employees, but that is usually it. The identity, much less the characteristics, practices, or time allocation of individual managers is rarely known. Furthermore, managerial inputs can be very abstract. It's not just time allocation that matters, but what the manager does with their time, like how they incentivize workers or deal with suppliers.

A recent set of papers has made considerable efforts to close this measurement gap. Some have focused on single-industry or even single-firm case studies by necessity, given the detail required in the data. More comprehensive efforts that cover a broader cross section of economic activity are underway, however.

Bloom and Van Reenen (2007)—hereafter BvR—offer one of the most comprehensive studies relating management practices (though not managers per se) to productivity. They and their team surveyed managers from over 700 medium-sized firms in the US, UK, France, and Germany. They surveyed plant managers, so the measured practices revolve around day-to-day and close-up operations rather than the broader strategic choices made at the executive level.

Surveys were conducted over the phone by a questioner who shared the respondent's native language. Information was probed on 18 specific management practices in four broad areas: operations, monitoring, targets, and incentives. The interviewers scored the firm on its practices based on these responses. Given the inherently subjective element of this measurement process, BvR took several steps to enhance accuracy and consistency. Managers were not told they were being scored. Questions on practices were open-ended (e.g., "Can you tell me how you promote your employees?" rather than "Do you promote your employees based on tenure?"). Financial performance was not discussed. The firms were small enough so that the interviewers would not already be aware of the performance of the firms they surveyed. Each interviewer conducted dozens of interviews, allowing BvR to control for interviewer fixed effects when relating management scores to outcomes. Further, over 60 firms were surveyed twice, by different interviewers; the correlation between the separate management practice scores for the same firms was 0.73.

Much of what was scored as "best practice" management in the interviews was based on the recommendations of the management consulting industry. This raises concerns about whether these practices are actually related to performance, or just the management fad of the moment. Importantly, therefore, BvR document that higher-quality management practices (and higher scores) are correlated with several measures of productivity and firm performance, including labor productivity, TFP, return on capital, Tobin's Q, sales growth, and the probability of survival.[9] The correlation between a firm's management practice score and its total factor productivity is statistically strong and economically nontrivial. Spanning the interquartile range of the management score distribution, for example, corresponds to a productivity change of between 3.2 and 7.5 percent. This is between 10 and 23 percent of TFP's 32 percent interquartile range in their sample.

BvR show two factors are important predictors of the quality of management practice in a firm. More intense competition in the firm's market, measured in several ways, is positively correlated with best-practice management. Additionally, management practice scores are lower when the firm is family-owned *and* primogeniture determined the current CEO's succession—i.e., he is the eldest son of the firm's founder. (I will discuss the competition-productivity link more extensively in Section 4. Interestingly, primogeniture's tie to productivity is not about family ownership *per se*—in fact, family ownership in isolation is positively correlated with good management.) These two factors are responsible for explaining most of the difference between the country-level average management scores in the sample. The variation in these averages is largely the result of the UK and France having a left tail of poorly managed firms. Both countries have traditionally favored primogeniture by tradition and family-firm exemptions in their inheritance tax laws.

Disentangling whether these correlations are causal is more challenging. Perhaps management consultancies base their recommendations on the practices observed at successful firms, but some excluded factor drives both management practice and performance. BvR, aware of this issue, estimated a specification in an earlier working paper version of the article that used competition and primogeniture measures to instrument for management scores. The notion is that the competitive and legal environments are orthogonal to other factors that drive management practices, at least in the short run. The estimated effect of management practices on TFP remains statistically significant and is in fact larger than the OLS case. This may suggest that unobserved third factors have a modest role, if any, and that BvR's management practice scores reflect (albeit noisily) true managerial acumen.

[9] The data from this paper is available on-line at http://cep.lse.ac.uk/_new/publications/abstract.asp?index=2313

BvR have since expanded their management practice survey program to gain greater coverage of business practices across economies. Bloom and Van Reenen (2010) and Christos Bloom, Genakos, Raffaella Sadun and Van Reenen (2010) review results from an extension of this survey program to nearly 6000 firms in 17 countries, including fast-growing China, India, and Brazil. The broader results echo those above. A particularly interesting pattern emerging from the early analysis is that the much lower average management practice scores in China, India, and Brazil are driven not so much by lower productivity across the board (though this is present to some extent), but in particular by a large left tail of very poorly managed firms. This has obvious implications for how trade growth and its assorted competitive pressures might impact productivity evolution in these and other countries. (More about Chinese and Indian firms' TFP levels below.) BvR are now further expanding the survey program to incorporate a panel element. This will be extremely useful, as it will allow one to control for unobservable fixed heterogeneity across firms as well as to see how firms' management practices change when their external environment does.

Other work in this vein includes James B. Bushnell and Catherine Wolfram (2009), who find that power plant operators have nontrivial impacts on the thermal efficiency of power plants. The best can boost their plant's fuel efficiency by over three percent, saving millions of dollars of fuel costs per year. Unfortunately, the data are less clear about what particular actions or attributes predict good plant management.

These research lines study managerial actions and policies at levels below the executive suite. Other work has focused on how those at the apexes of corporate hierarchies influence performance. Marianne Bertrand and Schoar (2003) study top executives (e.g., CEOs, CFOs, Presidents, etc.) who manage at least two firms for three years each during their 1969-1999 sample period. Following managers across multiple firms lets them test if individual executives can explain variation in firms' performance measures. While they don't measure productivity specifically, they do find that the individual manager fixed effects (particularly for CEOs) have significant explanatory power over firms' returns on assets. Adding these fixed effects to a regression of returns on firm and year fixed effects raises the adjusted R^2 from 0.72 to 0.77.

These results reflect performance differences that can be explained by the identity of the managers. This still leaves open the question of what the managers *do* or *know* that affects performance. Bertrand and Schoar don't have the sort of detailed management practice data of

BvR, but they do regress their estimated manager fixed effects on two variables they observe for the executives in their data: age and MBA attainment. They find that while age is not a significant factor, managers with MBAs have significantly higher ROA effects (by roughly 1 percent, as compared to a mean of 16 percent). This might be due to their more aggressive behavior as reflected in investment, leverage, and dividend-paying (or lack thereof) choices. More recent work (e.g., Steven N. Kaplan, Mark M. Klebanov, and Morten Sorensen (2009), Ulrike Malmendier and Geoffrey Tate (2009)) has started to dig deeper into how particular CEO practices and philosophies are tied to firm performance.

Other within-firm work has suggested that the human resources components of management, in particular, can affect productivity. This research—see for example Ichniowski, Shaw, and Giovanna Prennushi (1997); Lazear (2000); Barton H. Hamilton, Jack A. Nickerson, and Hideo Owan (2003); the papers cited in Ichniowski and Shaw (2003); Bruce Shearer (2004); and Bandiera, Barankay, and Rasul (2007 and 2009)—uses highly detailed, production-line-specific data to tie non-standard HR management practices like pay-for-performance schemes, work teams, cross-training, and routinized labor-management communication to productivity growth. These papers have elucidated some interesting details about the productivity effects of these practices. For instance, these practices may be complements: while they may have only modest impact on productivity when implemented in isolation, their total impact is larger than the sum of its parts when used in conjunction. Further, these practices are likely to have heterogeneous effects across production lines, even in the same plant, if different lines produce product variants of varying complexity. Brent Boning, Ichniowski, and Shaw (2007), for example, find an interaction between the complexity of the production process and the ability of innovative HR management in raising productivity.

Alexandre Mas (2008) shows in a vivid case study how poor management-labor relations can have productivity effects. He looks at the resale values of equipment made at plants and times where Caterpillar was experiencing labor strife during the 1990s. Compared to otherwise identical products made at plants or times without unrest, these products had about 5 percent lower resale values. This substantial productivity impact due to the implied reduction in the equipment's quality-adjusted service flows totaled $400 million.

With these and other studies, the evidence that management and productivity are related is starting to pile up. Further, some of this work strongly suggests that this relationship is causal.

Still, establishing causality definitively remains a key issue for research. Bloom, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2010) are attempting to establish as much by using what many consider to be the gold standard for establishing causality: a randomized field experiment. They are providing management consulting to a random set of Indian firms and will compare productivity growth in this treatment group to that observed in a set of control firms not receiving the intervention. This study could go a long way toward establishing whether or not a causal link exists. Any such link would raise additional questions. First, even if the interventions raised productivity, were they cost effective? That is, would they pay for themselves in a market setting? Second, given what we know about Indian firms in general, particularly for the left tail of the productivity distribution, if management consulting were to be effective anywhere, it would be in India. Should the experiment therefore be thought of as measuring the upper bound of the causal effect of management practices?

3.2. Higher-Quality General Labor and Capital Inputs

Management is an unmeasured input in most production functions, and hence is embodied in the productivity measure. Similarly, the productive effects of inputs like (non-management) labor and capital can also enter productivity if there are input quality differences that standard input measures do not capture.[10]

There is of course an enormous literature on human capital, far too large to cover here, that has tied several factors to labor quality, including education, training, overall experience, and tenure at a firm. Much of this work in labor economics has focused on wages as the outcome of interest. A smaller set of work has looked at human capital's impact on productivity.

Newer work using matched employer-employee datasets, which allow individual workers to be tracked across plants or firms over time, has offered evidence on the importance of labor quality. Abowd et al. (2005) offer a broad survey of the early evidence from these types of datasets, which tend to be newly constructed and therefore still have short panel histories. Their

[10] Attempts to capture labor quality differences in labor measures rather than productivity are the impetus behind using the wage bill to measure labor inputs rather than the number of employees or employee-hours. The notion is that market wages reflect variations in workers' contributions to production; firms with more productive workers will have a higher wage bill per employee. Of course, there are problems with this approach: wage variation might reflect the realities of local labor markets, or causation could be in the other direction, if more productive producers earn rents that are shared with or captured by employees (Van Reenen (1996); Abowd, Margolis, and Kramarz (1999)). Hence, more direct labor-quality measures are needed to definitively pin down labor quality's productivity contribution.

applicability for studying productivity, while limited now, will greatly increase over time. Still, some progress has been made with such data. Pekka Ilmakunnas, Mika Maliranta, and Jari Vainiomäki (2004), for example, use Finnish matched worker-plant data to show that (not surprisingly) productivity is increasing in workers' education as well as age.

As great a potential as such data may hold, the results in Fox and Smeets (2009) suggest that matched employer-employee data will not answer all of the literature's burning questions. They use matched employer-employee records from the Danish economy to control for worker education, gender, experience, and industry tenure in production function estimation. While these labor quality measures have significant coefficients in the production function, accounting for their influence only decreases the average within-industry 90-10 percentile productivity ratio from 3.74 to 3.36. There is plenty of productivity variation left to be explained. In a somewhat encouraging find for researchers using more limited data sets, they find that including the wage bill alone as a measure of labor inputs—data that is almost always available—does almost as well as including the full array of their human capital measures, though they caution that wage bills are subject to endogeneity concerns, as discussed above. This finding of only a modest role for finer labor skills measures in explaining productivity differences is echoed in Fernando Galindo-Rueda and Jonathan Haskel's (2005) investigation with similar U.K. data. (Incidentally, using the decline in productivity dispersion as a metric of a newly measured factor's importance in explaining productivity—or an R^2-type measure as Bertrand and Schoar use—is a good idea. Studies seeking to explain productivity dispersion should strive to conduct and report similar exercises.)

Capital can also vary in quality in ways not captured with standard measures. If capital vintages differ from one another in how much technological progress they embody, the common book-value-based capital stock measures will tend to miss variations in average capital vintages across producers. Several studies have tried to measure the rate of capital-embodied technological progress by carefully constructing measures of the distribution of capital vintages within plants or firms. Plutarchos Sakellaris and Daniel J. Wilson (2004) do exactly this using the annual investment histories of plants in the U.S. Annual Survey of Manufactures and industry-year-specific depreciation measures. They estimate a production function that is standard in all respects except that, rather than measuring capital inputs with single dollar-valued stock, they use a weighted sum of the plant's past investments. The weights combine the cumulative depreciation of a particular vintage's investment and a technological progress multiplier that they estimate. They assume that capital efficiency units grow at a constant rate per year, which they estimate to be between 8 to 17 percent per year, depending on the specification. These numbers are striking in their implications about how much productivity growth can come from investment alone. (Note that, unlike the standard capital deepening effects of investment that serve only to shift labor productivity, capital-embodied technological progress also raises TFP.) Other studies using different methodologies (e.g., Jason G. Cummins and Giovanni L. Violante 2002) have found somewhat smaller values, on the order of five percent per year. This seems to be an area desperate for further evidence, given its potential importance.

Van Biesebroeck (2003) measures the productivity impact of auto assembly plants shifting to "lean" technologies, which in that context involves new capital plus a host of complementary practices (teamwork, just-in-time ordering, etc.). This is also clearly related to the managerial practice discussion earlier. He finds that both the entry of new lean plants and the transformation of earlier vintage plants are responsible for the industry's acceleration of labor productivity growth during the late 1980s and early 1990s. Interestingly, his estimates of each technology's parameters suggest that capital-augmenting productivity is the primary driver of labor productivity growth under lean processes, while Hicks-neutral TFP-type productivity drives growth in the traditional technology plants.

Of course, not just physical capital can have unobservable quality differences. Certain types of capital may be themselves invisible—that is, intangible capital. Such capital can include any of a number of concepts, like a firm's reputation, know-how, or its loyal customer base, just to name a few. Despite the difficulty in quantifying these types of capital, they can have very real output effects that, as such, will result in measured productivity differences. I will discuss some specific cases of intangible capital in operation below, but the full breadth and depth of intangibles' role in explaining productivity differences are still very much open questions.

3.3. IT and R&D

While the research described above indicates that input heterogeneity matters, the productivity effects of a particular type of capital—information technology (IT)—have been the subject of intense study. This is rightly so; many have hypothesized that IT was behind the resurgence in U.S. aggregate productivity growth in the mid-1990s after 20 years of sluggish performance, and that IT has more generally influenced productivity patterns across multiple industries and countries. Given the sheer size of GDP per capita variation that can be driven by even a modest change in trend productivity growth over a sustained period, it is not surprising that sources of such changes receive considerable research attention. Because of this attention, I discuss the work done on this particular capital type separately here.

An overview of IT capital's broad productivity impacts, particularly in driving the growth resurgence, can be found in Dale W. Jorgenson, Mun S. Ho, and Kevin J. Stiroh (2005, 2008) and Stephen D. Oliner, Daniel Sichel, and Stiroh (2007). These studies document that IT-related productivity gains—both spectacular productivity growth in IT-*producing* industries and more modest changes in IT-*using* industries—play an important role in explaining aggregate U.S. productivity growth over the past couple of decades.

At the same time, Bart van Ark, Mary O'Mahony, and Marcel P. Timmer (2008) show that the EU's comparably sluggish productivity growth over the same period can be explained in large part by the later emergence and smaller size of IT investment in European economies. Bloom, Sadun, and Van Reenen (2009a) suggest that it is not geography *per se* that matters, but rather the location of the owning firm. They show U.S.-based multinationals operating in the EU are more productive than their EU counterparts, and this productivity advantage is primarily derived from IT capital. They link their management practices data discussed above to data on IT usage to test for particular mechanisms through which this productivity advantage arises. Their evidence points to a complementarity between IT capital and human resources practices, explaining U.S. multinationals' productivity advantage in the EU.

These broad patterns raise the question of which specific micro mechanisms actually underlie the aggregate relationship between IT and productivity growth. Several studies have explored this issue with detailed production data. Thomas N. Hubbard (2003) shows how on-board computers raise average utilization rates of trucks that they are installed in. The computers provide dispatchers real-time information on a truck's locations and load status, allowing them to better match the available cartage capacity to innovations in demand.[11]

[11] Adopting any new technology, IT or otherwise, obviously has its own costs. A new technology's net productivity benefit to the adopter depends on the difference between the increased production the new technology facilitates and its acquisition cost. For the marginal adopting producer, this net gain will be zero. However, inframarginal producers experience positive productivity gains. The aggregate productivity gains that any technology will offer

Ann Bartel, Ichniowski, and Shaw (2007) show how better computer numerically controlled (CNC) machining centers—automated devices that shape parts from raw material stock—raise productivity in the valve manufacturing industry by shortening setup times, raising speeds of production runs, and even allowing quicker inspections. The appealing element of the study's empirical approach is that both the products and the production process, except for the particular pieces of IT capital whose contribution is of interest, remain constant across observations. The paper also shows that IT-intensive product design tools like computer-aided design packages make it easier to design customized parts, and lower setup times make multiple production runs less costly. Offering a broader array of parts allows the firms to better match their production capabilities to their customers' desires, increasing the surplus of their sales.

Such a gain in surplus from product specialization raises an important broader point about productivity measurement. Better customization from IT can raise firms average product prices. Measures of productivity in physical units of output (e.g., number of valves per unit input) may therefore not fully capture the surplus gained. This is one case where the limit of most producer-level data sets to revenue-based output measures does not pose a measurement problem, because this sort of productivity gain would be reflected in revenues but not physical quantities. (That said, the concern about price variations due to local market power or demand shocks creating productivity mismeasurement still applies in differentiated product settings.)

Erik Brynjolfsson, Andrew McAfee, Michael Sorell, and Feng Zhu (2007); Bartelsman, Pieter Gautier, and Joris de Wind (2009); and Giulia Faggio, Kjell G. Salvanes, and Van Reenen (2009) each draw, in related but distinct ways, broader lines connecting IT and productivity. Brynjolfsson et al. document case studies where IT enhances the speed with which firms can replicate practices they find productive in one of their lines of business across the entire organization. This ability to lever-up a productivity advantage means successfully innovating firms displace less productive competitors more quickly. IT thus raises the volatility of firm performance. Brynjolfsson et al. test for and find this heightened volatility in a sample of Compustat firms in 61 industries. In the context of the model in Section 2, Brynjolfsson et al. essentially argue that IT reduces the concavity of the firm's revenue function, allowing them to

will therefore also depend on the competitiveness of the technology-producing sector. A lower markup and price for the technology raises both the number of inframarginal adopters and the net productivity gain that each experiences.

better leverage (and in a dynamic world, do so more quickly) any inherent productivity advantages (increases in A_i) that they develop or stumble upon.

Bartelsman, Gautier, and de Wind (2009) further develop the notion that IT shifts not just the mean of the distribution of innovation outcomes but its variance as well. Because poor outcomes are truncated by the option to exit—again in the parlance of the model above, firms drawing a productivity level below $\underline{A}$ don't need to produce at a loss—greater variance raises the value of making risky innovations. Bartelsman et al. note, however, that exit costs (absent in the model in Section 2) will stifle firms' willingness to innovate because they make it harder to dismiss unsuccessful outcomes. They argue that employment-protection legislation like firing costs makes exit more expensive and therefore reduces firms' willingness to adopt IT. They show that IT-intensive sectors are in fact smaller in countries with greater legal restrictions on firms' abilities to close unsuccessful lines of business. They cite employment protection legislation as a major contributor to the IT gap documented by van Ark, O'Mahony, and Timmer (2008). (I will further discuss the role of flexibility in input markets further in Section 4 below.)

Faggio, Salvanes, and Van Reenen (2009) document that within-industry productivity dispersion in the UK has trended upwards over the past couple of decades. They relate this increased dispersion to the growth in wage dispersion that has occurred over the same period in the UK and almost every other developed economy. It would be interesting to see if similar productivity spreading is occurring in concert with wage dispersion growth in these other economies. More directly applicable to the theme of this section, however, is that Faggio et al. show that industries which experienced the greatest growth in productivity dispersion also saw the largest increases in IT capital intensity. This is yet more evidence tying IT to greater productivity variance.

There is a long literature linking R&D and productivity, and recent additions to it have focused on exploring the ties at the micro level. As with many input-based stories of productivity differences, the difficulty is in separating correlation from causation. There are many reasons why more productive firms might do more R&D, suggesting that some of the causation may go the other way.

Ulrich Doraszelski and Jordi Jaumandreu (2009) model firm productivity growth as the consequence of R&D expenditures with uncertain outcomes. Estimating their model using a panel of Spanish firms, they find that R&D does appear to explain a substantial amount of productivity growth. However, and picking up the theme of increased variance tied to IT capital discussed above, they also find that firm-level uncertainty in the outcome of R&D is considerable, much more so than with respect to the return on physical capital investment. In fact, their estimates suggest that engaging in R&D roughly doubles the degree of uncertainty in the evolution of a producer's productivity level.

Bee Yan Aw, Mark J. Roberts, and Daniel Yi Xu (forthcoming) highlight the bidirectional causality between R&D and productivity in their study of Taiwanese electronics exporters. They find that firms that select into exporting tend to already be more productive than their domestic counterparts (more on this in the trade section below), but the decision to export is often accompanied by large R&D investments. These investments raise exporters' productivity levels further in turn, highlighting both selection and causal effects tying productivity to R&D. The timing of this R&D blitz is consistent with a world where the exporters are more willing to innovate on the margin because they can spread the potential gains of productivity growth across a larger market.

Of course, R&D is simply one of the more observable components of firms' overall innovative efforts. Many firms undertake both process and product innovation without formally reporting R&D spending. (I will discuss product innovation's ties to productivity differences in further detail below.) This limits the literature's ability to give a comprehensive look into the relationships between productivity and innovation. Still, it is a very useful start, and the mechanisms the R&D literature highlights are likely to often overlap with the effects of unmeasured innovative spending.

3.4. Learning-by-doing

The very act of operating can increase productivity. Experience allows producers to identify opportunities for process improvements. This productivity growth, often called learning-by-doing, has a long and rich history of study in the literature, but has recently been investigated in more detail given newly available micro-level production data.

Benkard (2000) studies the precipitous drop in the labor hours Lockheed needed to assemble its L-1011 TriStar wide-body aircraft. The first few units off the line required more than one million person hours (equivalent to three shifts a day of 2500 workers each for fifty work days). This was cut in half by the 30th plane, and halved again by the 100th. Benkard estimates both the learning rate—how fast past production increases productivity (decreases unit labor requirements)—and the “forgetting” rate, which is how fast the knowledge stock built by learning depreciates. Forgetting is quantitatively important in this setting: Benkard estimates that almost 40 percent of the knowledge stock depreciates each year. This may not be literal forgetting, but could instead primarily reflect labor turnover. An additional factor in “forgetting” was the shift to a new variant of the plane after about 130 units. This new variant was different enough that the imperfect substitutability of the knowledge stock between the original and new variants led to a temporary but substantial increase in labor requirements.

Rebecca Achee Thornton and Peter Thompson (2001) investigate what *types* of experience matter in productivity growth from learning by doing. Their data includes unit labor requirements for several design variants of 4000 Liberty ships produced by multiple shipyards during World War II. The multi-design/multi-yard nature of the data lets them estimate the relative productivity contributions of four different measures of past production experience: the yard’s past production experience with a particular design, the same yard’s past production of other designs, other yards’ experience with the particular design, and other yards’ production of other designs. Not surprisingly, a yard’s past production of a particular model matters most for productivity growth in that same model. After that comes the yard’s experience with other ship designs, at about 60 percent the size of the own-design effect. Cross-yard spillovers are considerably smaller—only about five to ten percent of the own-yard, own-design learning impact. These cross-plant learning effects, while relatively modest here, do show that producers may become more productive by learning from other businesses. I will discuss cross-business spillovers more below.

In Steven D. Levitt, John A. List, and Syverson (2010), we find more limited cross-model learning spillovers within an auto assembly plant. Using detailed data on hundreds of individual operations during assembly of thousands of cars, we studied the causes and effects of manufacturing defects. This particular plant began production of three model variants (nameplates) of a common platform at staggered times during a production year. Each time a new model ramped up, the plant began a new learning curve with defect rates roughly as high as during the previous model’s ramp-up. There is some evidence, however, that learning happens faster for the later models: defect rates fall to their long-run level more quickly. An interesting contrast was seen when looking at what happened to defect rates when a new shift started producing a given model. In that case, re-learning was not necessary. The new shift began operating at defect rates at about the same level as the previous shift had achieved after it already had run down much of the learning curve.

Ryan Kellogg (2009) looks at oil and gas drilling in Texas to study how learning occurs when an upstream and downstream producer work together over time. He follows the efforts of pairs of producers and drillers. The former are companies actively involved in exploring for, extracting, and selling oil, while the latter firms specialize in boring out the wells that the producers hope will yield oil. Since producers typically work with multiple drillers and vice versa, and work in different fields, Kellogg is able to separately measure the productivity impacts of the experience of producers alone (i.e., regardless of the drilling firms they work with), drillers alone, and the joint experience of producer-driller pairs. He finds that accumulated experience between a producer-driller pair increases productivity above and beyond that of each of the firms' overall experience levels. This relationship-specific experience is a type of capital that is lost if the firms split up, giving them incentives to preserve their contracting environment.

3.5. Product Innovation

Innovations in product quality may not necessarily raise the quantity of output (measured in some physical unit) per unit input, but they can increase the product price and therefore the firm's revenue per unit input. If one thinks about productivity as units of quality delivered per unit input, product innovation can enhance productivity. This is captured in standard revenue-based productivity measures since they reflect price variations across an industry's plants or firms. (Though as mentioned above and discussed further below, revenue productivity can also be misleading when price variation due to differences in market power across producers exist.) Product innovation can be aimed at entering new markets, or at refocusing a firm's efforts toward growing demand segments, as documented in Daron Acemoglu and Joshua Linn (2004).

Product innovation's productivity effects have been studied in several recent papers. As touched on above, one of the mechanisms behind IT-based productivity growth that Bartel, Ichniowski, and Shaw (2007) point to is an improved ability to customize products. Other inputs mentioned above, like R&D and higher-quality employees, can also spur innovation.

Rasmus Lentz and Dale T. Mortensen (2008) use Danish firm-level data to estimate a model of firms' product innovation efforts in the vertical-quality-ladder style of Tor Jakob Klette and Kortum (2004). They find that about 75 percent of aggregate productivity growth comes from reallocation of inputs (employment in their setup) to innovating firms. About one-third of this comes from entry and exit channels. The other two-thirds occurs as inputs move toward growing firms (and hence innovating firms, as seen through the lens of their model) from firms that lose market share when they fall behind the quality frontier.

Natarajan Balasubramanian and Jagadeesh Sivadasan (forthcoming) link detailed and broad-based data on firms' patenting and production activities (they merge the NBER patent database with the U.S. Census Business Register) to see what happens when a firm patents. They find clear evidence that new patent grants are associated with increases in firm size (by any one of a number of measures), scope (the number of products it makes), and TFP (though the evidence is weaker here). Whether these correlations reflect the causal effects of patents is not clear; patenting activity could be just one part of a firm's coordinated push into new markets. Nevertheless, given the breadth of the study's coverage and its result that correlations exist, more research in this area would be worthwhile.

Bernard, Stephen Redding, and Peter Schott (2010) show that a firm's TFP is positively correlated with the number of products it produces. This holds both in the cross section and within firms over time. At the very least, these results indicate that productivity growth accompanies expansion of the variety of products a firm offers. It is less clear whether innovative activity drives both productivity and product-variety growth, or whether firms experiencing general productivity shocks "strike while the iron is hot," expanding their product offerings in response. The role of changes in product scope in firm size and productivity growth is one that is just beginning to get the attention it deserves in research agendas.

3.6. Firm Structure Decisions

A lot of the micro productivity literature uses the establishment (e.g., factory, store, or office) as the unit of analysis. This is in part data driven; many surveys are conducted at this level. Plus, plants often embody the smallest indivisible unit of a production process, and as such are a natural level at which to study technologies. But it is also clear that firm-level factors, and in particular, the organizational structure of the firm's production units—the industries they operate in, their vertical and horizontal linkages, their relative sizes, and so on—will sometimes be related to the productivity levels of the firm's component business units.

Some have suggested there is a link between firm decentralization and how easily productive new technologies are adopted. Bloom, Sadun, and Van Reenen (2009b) favor this explanation for European firms' recent laggard productivity growth (as mentioned above). It is also the subject of Acemoglu, Philippe Aghion, Claire Lelarge, Van Reenen, and Fabrizio Zilibotti (2007). The evidence tends to be suggestive but indirect, however, and this is an area where careful work in measuring firm structures (not an easy task) could pay big dividends.

Silke Forbes and Mara Lederman (2009) look at how vertical integration affects airline performance. They find that, among flights departing from a given airport on a given day, airlines that own their regional affiliates experience shorter delays and fewer cancellations than those contracting with affiliated regionals at arms' length. This performance advantage appears to come largely from differential performance on adverse weather days. Forbes and Lederman posit that contracts are limited in their ability to fully specify contingent actions necessary to react most effectively to short-horizon logistical problems. Vertical integration, by clearly setting out the decision rights within the organization, allows airlines to more nimbly respond to unexpected scheduling issues. This flexibility comes at a cost, however: primarily in higher wage costs for integrated airlines. This could explain why not every mainline carrier has integrated.

Hortaçsu and Syverson (2009) use the Longitudinal Business Database, which contains most private non-agricultural establishments in the U.S., to examine the productivity of plants in vertically structured firms. We find that vertically integrated plants have higher productivity levels than their non-integrated industry cohorts, but most of this difference reflects selection of high-productivity plants into vertical structures rather than a causal impact of integration on productivity. Surprisingly, these productivity differences—and indeed the firm's choice to have a vertical structure at all—usually are not related to the movements of goods along the production chain. Vertically integrated firms' upstream plants ship a surprisingly small amount to downstream plants in their firm (small relative to both the firms' total upstream production and their downstream needs). Roughly one-third of upstream plants report no shipments to their firms' downstream units; half ship less than three percent of their output internally. This suggests that rather than moderating goods transfers along production chains, integration instead allows more efficient transfers of intangible inputs (e.g., managerial oversight) within the firm.

Maksimovic and Phillips (2002) and Schoar (2002) both investigate the productivity of plants within conglomerate firms (in their setting, those that operate in multiple two- or three-digit SIC industries). Their work was spurred on in part by the extensive finance literature on the "diversification discount," the term for the oft-measured negative correlation between a firm's financial returns and the number of business lines it operates. Both papers leverage U.S. manufacturer microdata to convincingly argue that the diversification discount is not about low productivity (or even, in one case, any sort of underperformance). They differ, however, in their explanations.

Maksimovic and Phillips (2002) make a selection argument. Firms that choose to specialize are likely to have idiosyncratically high productivity draws in a particular line of business, but considerably weaker draws outside this segment. Firms that choose conglomerate structures, on the other hand, are likely to have high draws in several industries, but not exceptionally high draws in any particular industry. Thus if one simply compares the performance of a conglomerate's segments to the focused and highly productive segments of a specialist, the latter would expectedly be higher. This result does not rely on the previous literature's favored explanations of management overreach, cross subsidization of weak segments, or other agency problems at conglomerates. It simply reflects the optimal allocation of resources within a business given the firm's inherent abilities. They support their efficient allocation argument by showing that conglomerate firms' most productive plants are in their largest segments, and segments of a given rank are more productive in larger firms. Furthermore, conglomerates expand on their strongest margins: their largest, most productive segments are more sensitive to demand shifts than their smaller, less efficient lines of business.

Schoar (2002) notes that in her sample, plants in conglomerates have, if anything, higher permanent productivity levels. The observed discount reflects the temporary adjustment costs resulting from the very act of diversifying into new businesses. She shows that when a conglomerate diversifies, the plants it buys actually experience productivity growth, suggesting that they are in fact being reallocated to more capable management (there will be more on the reallocation of productive inputs below). At the same time, however, the conglomerate's existing plants suffer productivity losses. Since conglomerates have on average many more existing plants than acquired ones, average productivity in the firm falls for a period. Schoar attributes these productivity changes to a "new toy" effect: managers (over-) concentrate their efforts on integrating the new plants and business lines at the expense of existing ones. She also finds evidence that the firms' wages absorb any performance rents, also leading to a bifurcation between performance as measured by productivity and by stock market returns.

4. External Drivers of Productivity Differences

The previous section discussed factors that operate within the firm to determine productivity levels. Producers have, at least in theory, some degree of control over these factors. This section focuses instead on how producers' operating environments can influence productivity levels and growth. These environmental factors may not operate directly on productivity, but they can affect producers' incentives to apply the factors discussed in the previous section. They can also influence the extent to which such efforts are successful at moving producers to a higher position within their industry's productivity distribution, and how responsive market share and survival are to productivity differences. That is, these external drivers can impact both the so-called "within" and "between" components of aggregate productivity growth. The within component comes from individual producers becoming more efficient. The between component arises when more efficient producers grow faster than less efficient ones, or when more efficient entrants replace less efficient exiting businesses.[12]

By their nature, these environmental elements are also the most closely tied to government policy. Therefore understanding these drivers merits special attention when considering the productivity implications of market interventions.

Before discussing the specific external drivers, it is worth taking a minute to discuss why the operating environment should affect observed productivity levels. The most basic producer theory, after all, says any profit-minimizing firm minimizes its cost of producing its chosen quantity. This prediction is invariant to the structure of the market in which the firm operates.

[12] Many studies attempt to quantify the relative contributions of within and between effects by decomposing aggregate productivity growth into terms that reflect the separate effects. Petrin and Levinsohn (2008) have recently raised caveats about the robustness of these commonly used "accounting decompositions." They advocate a method that focuses on measuring the gaps between the estimated social marginal benefits and costs of producers' inputs. Aggregate productivity grows when inputs are reallocated in a way that reduces the average gap. While distinct in theory and empirical implementation from the accounting decompositions, such "gap methods" have the same conceptual goal: to separately measure how much aggregate productivity growth comes from businesses becoming more efficient themselves and how much comes from reallocation of economic activity to more efficient producers.

The presence of spillovers is one possible channel through which the external environment affects productivity levels. I discuss situations where other firms' production practices influence another business's productivity level first in this section.

A second possibility is that external drivers influence the extent of Darwinian selection in the firm's market. This force is highlighted by the model in Section 2. Environmental factors that shift the model's exogenous parameters or the shape of the revenue function will change the minimum productivity level necessary for profitable operation, $\underline{A}$, and the responsiveness of market share to productivity differences. This will shift the observed productivity distribution among the market's producers.

Even in the absence of spillovers or selection, external factors can affect producers' incentives to raise their own productivity level. How can this be if theory says firms minimize costs? Well, the standard, static cost-minimizing firm model is an inadequate description of the technology adoption processes. A richer model like that in Thomas J. Holmes, David K. Levine, and Schmitz (2008)—who build off Kenneth Arrow's (1962) seminal work—points out additional channels through which a firm's market environment (and the competitive structure in particular) shifts producers' incentives to raise their productivity level. Holmes et al. suppose that adopting a productivity-enhancing practice involves disruption costs: a temporary period where costs are actually *higher* than before any technological change was made. Disruption could be due to installation issues, fine-tuning new technology, retraining workers, and so on. With such adoption costs, producers facing less competition have less incentive to adopt the new technology, because the higher per-unit profits that monopoly power brings raise the opportunity cost of changing production practices. In the context of the model in Section 2, less competition means a more concave revenue function due to steeper residual demand curves. This could arise from, for example, less scope for consumers to substitute across producers in the output market.[13]

The reality of production is also much more complex than even in these augmented models. Most technologies, even if routinized, are intricate, multifaceted processes that require

[13] A second, more roundabout mechanism also relates greater competition to technology innovation and adoption. If heightened competition raises the firm's probability of exit or bankruptcy, the convexity of the firm's payoffs created by limited liability encourages risk-taking (see, for example, Susan Rose-Ackerman (1991)). In essence, competition may drive desperate firms to "throw a Hail Mary" by adopting risky but potentially productive new technologies in the hope of staving off collapse. I will discuss another implication of the convexity of firm payoffs and technology adoption below.

considerable coordination. They are constantly being buffeted by shocks to input costs and demand-driven shifts in capacity requirements. Cost-minimizing production practice is really therefore a moving target, a constantly shifting ideal combination of operations. Elements of a firm's market environment can affect the firm's incentives to chase that moving target.

4.1. Productivity Spillovers

Producer practices can have spillover effects on the productivity levels of other firms. These externalities are often discussed in the context of classic agglomeration mechanisms like thick-input-market effects and knowledge transfers. Knowledge transfers in particular need not be tied to any single geographic or input market. Producers are likely to attempt to emulate productivity leaders in their own and closely related industries, regardless of whether they share a common input market.

Any empirical search for spillovers must face the classic "reflection problem" familiar to the peer effects literature: correlated productivity levels among cohorts of producers can be a sign of spillovers, but the correlation might also reflect the impact of common shocks from unobserved third factors. Obviously, if one can observe exogenous productivity shocks for a subset of producers and track how related producers' productivity levels evolve in response, this goes a great way towards identifying causality. Such instances can be difficult to observe generally, however, and such an approach cannot be used in a single cross section. An alternative strategy is to test whether the intensity of the productivity correlation is related to some measure of between-producer distance, be it in geographic, technological, or product-market space. Higher productivity correlations among "nearby" producers are predicted by many theories of spillovers. This approach is still imperfect, however, as the structure of common shocks might also be related to distance.

Enrico Moretti (2004) explores agglomeration-type productivity spillovers by matching the 1980 and 1990 U.S. Population Censuses with the 1982 and 1992 Census of Manufactures by city-industry. He estimates a plant-level production function that includes the share of workers in *other* industries in the metro area who have completed some college. He interprets the estimated marginal product of this outside educated labor as a productivity spillover. Moretti finds that the marginal product of the local human capital measure is in fact positive. Furthermore, the measured spillovers are stronger across plants that are "close" in both the geographic and technological senses. These results are consistent with both the thick-input-market and knowledge-transfer stories of productivity spillovers.

Several studies have focused specifically on the role of knowledge transfers. On one level, it seems that they must exist. It is doubtful that productivity-enhancing practices are completely excludable; businesses cannot always keep every facet of their production process secret. On the other hand, the ubiquity of large and persistent productivity differences within industries suggests that any such emulation/spillover process is far from perfect. Frictions clearly exist that prevent less efficient producers from fully replicating industry leaders' best practices. The crucial research questions of these studies, then, are the size of knowledge transfers, what features influence this size, and the channels through which the spillovers operate.

Rachel Griffith, Rupert Harrison, and Van Reenen (2007) show that the geographic location of a firm's R&D activity matters. Using patent data to pin down the historical locations of U.K. firms' R&D operations (they use pre-sample locations to minimize endogeneity of the location of research activity), they find that U.K. firms with a greater R&D presence in the U.S. have faster overall productivity growth, and that this growth is more highly correlated with the growth of the U.S. R&D stock in the same industry. These patterns are consistent with a U.S. research presence making it easier for firms to tap into the knowledge base of the U.S. economy, which tends to be the technological leader in most industries. The precise mechanism through which this technology tapping occurs is unclear, and would be an interesting area for further exploration.

Bloom, Schankerman, and Van Reenen (2007) point out that spillovers can cut two ways: technological spillovers can benefit everyone, but there can also be market-stealing effects on the product market side.[14] They separately identify these two effects by comparing the impact of firms' R&D (instrumented for using federal- and state-level R&D tax incentives) on other firms at varying technological and product market distances. They measure technological distance using correlations in firms' patenting patterns and product market distance using the correlation in firms' sales across business segments. Because these two distances are not perfectly correlated across firms, they can measure the separate impacts of R&D. They find that both

[14] Hans Gersbach and Armin Schmutzler (2003) demonstrate how product market competition can endogenously determine the extent of knowledge spillovers via labor mobility.

types of spillovers matter, but technological spillovers quantitatively dominate, creating a net positive externality.

Bartelsman, Haskel, and Ralf Martin (2008) make an interesting distinction between global and economy-specific technology frontiers. They show using microdata from numerous countries that a plant's productivity converges faster toward the productivity level of the domestic leader in its industry than the global industry leader. A second intriguing result is that if a plant falls sufficiently behind the global frontier, any pull toward convergence disappears, but convergence to the national frontier remains no matter the size of the gap (conditional on survival, of course).

Gustavo Crespi, Chiara Criscuolo, Haskel, and Matthew Slaughter (2008) and Wolfgang Keller and Stephen R. Yeaple (2009) also look at cross-border productivity convergence. Crespi et al. focus on measuring the information flows that could be the source of this convergence. They combine production microdata with survey data on where firms gather information used in their innovative efforts. They find that, not surprisingly, "nearby" firms (e.g., suppliers and competitors, though less so buyers) are primary sources; that much of this information, particularly from competitors, is free (though surely not given *freely*); and that having a multinational presence aids these flows. Keller and Yeaple (2009) tie productivity growth among publicly traded U.S. firms to foreign direct investment in those firms sectors by foreign-owned multinationals. FDI-driven spillovers account for a substantial portion of productivity growth, especially in high-tech sectors.

These papers and others like them suggest that spillovers exist and operate through various mechanisms, though again the observed productivity dispersion also makes clear that substantial frictions to the diffusion and replication of best practices remain. Policies meant to increase such spillovers must be careful, however, to not destroy firms' incentives to innovate. If spillover-enhancing policies make it too hard for firms to appropriate the benefits of their innovation, the policies could do more damage than good in the long run.

4.2. Competition

Pressures from threatened or actual competitors can affect productivity levels within an industry. Competition drives productivity through two key mechanisms; this section discusses examples of research into both.

The first is Darwinian selection among producers with heterogeneous productivity levels. Competition moves market share toward more efficient (i.e., lower-cost and generally therefore lower-price) producers, shrinking relatively high-cost firms/plants, sometimes forcing their exit, and opening up room for more efficient producers. It also raises the productivity bar that any potential entrant must meet to successfully enter. In the static model of Section 2, these mechanisms are summarized as an increase in $\underline{A}$. Such selection underlies the "between" component of aggregate productivity growth mentioned earlier.

The second mechanism acts through efficiency increases within plants or firms. As discussed above, heightened competition can induce firms to take costly productivity-raising actions that they may otherwise not. Besides raising producers' own productivity levels, this effect of competition leads to aggregate productivity growth via the "within" component. There is a Schumpeterian caveat to this within-effect of competition, however. As Xavier Vives (2007) points out, under certain conditions, heightened competition (at least for a market of fixed size) can actually diminish a firm's incentives to make productivity-enhancing investments.

Because of the substantial literature built around the productivity impacts of trade competition, I discuss it in a separate subsection below. I first cover general competitive effects.

4.2.1. Intra-market Competition

A general indicator that product-market competition is enhancing productivity is a positive correlation between productivity and producer growth and survival. Such correlations have been a robust finding in the literature; Foster, Haltiwanger, and Krizan (2001) offer a broad-based overview, for example.[15] Several recent studies have looked at particular mechanisms through which competition leads to a Darwinian selection process.

Syverson (2004b) investigates the connection between competition and productivity in a case study of the ready-mixed concrete industry, which is well suited for this type of investigation. The industry's physically homogeneous product and very high transport costs make spatial differentiation paramount. Differences in competitiveness across markets should

[15] Foster, Haltiwanger, and Syverson (2008) point out that these results linking selection to productivity actually reflect selection on *profitability*, since intra-industry price variation caused by idiosyncratic demand differences across plants is buried in standard revenue-based productivity measures. They show that such demand variation is extremely important in explaining plant survival patterns, even in their sample of plants in homogeneous-product industries. This broader interpretation of the evidence to include demand-side factors will be discussed further in the following section.

therefore be related to the density of concrete producers in the market. It is harder for inefficient concrete producers to be profitable in dense markets, because if they charge the high prices necessary to cover their costs, customers can easily shift to their more efficient competitors. This implies the productivity distribution of ready-mixed plants will be truncated from below as density rises. This is indeed what holds in the data. Markets with denser construction activity have higher lower-bound productivity levels, higher average productivity, and less productivity dispersion. (Construction density is used as an exogenous shifter of concrete producer density because the construction sector buys almost all of the ready-mixed industry's output, yet concrete accounts for only a small share of construction costs.) In Syverson (2007) I show that these patterns of competition-driven selection on costs are also reflected in ready-mixed prices.[16]

Outside of manufacturing, Foster, Haltiwanger, and Krizan (2006) find that aggregate productivity growth in the U.S. retail sector is almost exclusively through the exit of less efficient single-store firms and by their replacement with more efficient national chain store affiliates. This evokes stories surrounding the growth and competitive impacts of discount retailers like Wal-Mart and Target over the past two decades.

These studies focus on the selection effect of competition. Schmitz (2005) offers an example of productivity growth in an industry that is driven almost entirely by within-effect efficiency improvements. He follows U.S. iron ore mining during the period the industry was first facing competition from foreign producers. (Brazilian mines, specifically. We will discuss more examples of trade-induced productivity change in a separate section below.) The case study shows how competition can drive existing firms to improve their productivity.

The U.S. iron ore industry had been protected from foreign competition by the high costs of transporting ore from its other sources on the globe (e.g., Australia and Brazil). By 1980, however, increased production from low-cost Brazilian mines brought delivered prices for Brazilian ore in the Great Lakes region in line with delivered prices from northern Minnesota's Mesabi Range, the major ore-producing area of the U.S. Facing competition from abroad for the first time, the U.S. producers attempted to lower costs by making drastic changes in their production operations. Schmitz shows most of these changes centered on loosening the strict

[16] Such price effects also raise an interesting point given the common use of revenue-based productivity measures. Namely, as competition raises the average physical (i.e., quantity-, not revenue-based) productivity level in the market, it also reduces prices. This means standard revenue-based productivity measures will understate the true impact of competition on average productivity levels.

work rules in the U.S. mines. For instance, mine managers originally had very little flexibility in their ability to assign different workers to different tasks. The initiation of serious competition allowed the mines to gain back flexibility in new contracts, raising their utilization of available labor and enabling them to shed unneeded overhead workers. The reconfigured contracts were extremely successful at raising productivity. The industry's average labor productivity had been roughly constant at two tons of ore per worker-hour for several decades preceding 1980. By 1985, however, it had doubled to four tons per hour. As a result, the mines were able to remain competitive even in the face of continuing falling Brazilian ore prices.

Other recent studies have shown these detailed case studies appear emblematic of much broader competitive effects that act across numerous industries and economies. For example, Syverson (2004a) looks at the entire U.S. manufacturing sector. Richard Disney, Haskel, and Ylva Heden (2003a and 2003b) and the studies described in UK Office of Fair Trading (2007) show similar results in the U.K. And Giuseppe Nicoletti and Scarpetta (2005) overview evidence across OECD countries.

4.2.2. Trade Competition

As seen in Schmitz's results for the iron ore industry, the presence—or even just the threat—of imports from abroad is another form of competitive pressure. This phenomenon is the focus of a burgeoning line of research, driven in part by the recent theoretical trade literature focusing on heterogeneous-productivity producers and their response to trade, especially Eaton and Kortum (2002) and Melitz (2003).

Pavcnik (2002) shows how trade liberalization during the 1970s drove productivity growth among Chilean manufacturers. The paper demonstrates that sectors facing new import competition saw faster productivity growth over her 1979-86 sample period than sectors producing primarily non-tradables. Pavcnik goes on to show that these industry-level gains came both from existing producers raising their productivity levels (the within effect) and from the reallocation of activity away from—and sometimes, the exit of—less efficient, formerly protected producers (the selection effect).

Bloom, Mirko Draca, and Van Reenen (2009) look at how Chinese import competition—the proverbial 800-pound gorilla in trade policy discussions—affected productivity and innovation in 12 European countries between 1996 and 2007. To identify competition's effects, they exploit the differential across-product drops in import barriers that occurred when China became part (due to its accession into the WTO) of the now-expired Multi Fibre Agreement in 2001. European firms producing the products which saw the greatest increase in competition responded in one of two ways. Some, particularly those using low-tech production methods, shrank and exited. Others, however, innovated. Their patent rates, R&D, IT adoption, and TFP growth increased concurrently. In aggregate, therefore, Chinese trade competition increased aggregate TFP in these markets through both within- and between-firm (selection) effects.

Multiple studies using producer microdata have found comparable results in other settings. Examples include Marcela Eslava, Haltiwanger, Adriana Kugler and Maurice Kugler (2004); Muendler (2004); Bernard, Jensen, and Schott (2006); Ana M. Fernandes (2007); and Verhoogen (2008). The specific mechanisms through which trade-oriented competition is postulated to increase productivity do vary across the papers, from quality upgrading within plants to heighted selection across plants. Mary Amiti and Jozef Konings (2007) highlight a separate mechanism through which trade can increase productivity: the expansion of the set (or declines in the effective price) of intermediate inputs when imported inputs become more available. I will discuss the input-market channel further below.

Interestingly, despite the strong correlation between the average productivity level of an industry's plants and that industry's trade exposure, there is less evidence of large productivity impacts on the domestic plants when they begin exporting. That is, exporters are almost inevitably more productive than their non-exporting industry counterparts, but most studies have found that this correlation largely reflects selection rather than a causal impact of exporting on productivity. Plants that choose to begin exporting were *already* more productive before trade. This is surprising if only because exporting firms can leverage the benefits of any productivity gains across larger markets, raising their incentive to engage in innovative activities.

That said, Van Biesebroeck (2005) and Jan De Loecker (2007) document cases where exporters' productivity advantage grows *after* entry into the export market. (This is sometimes referred to as the "learning-by-exporting" hypothesis.) Both are in somewhat special settings, which might explain in part why they find post-export productivity growth while many others have not. The post-export growth of Van Biesebroeck's (2005) sample of Subsaharan African exporters appears to reduce their credit and contract enforcement constraints, allowing them to undertake what were previously prohibitively costly productivity-raising activities. Such a mechanism raises the question of whether it would apply to any firm that chooses to export (if so, why wouldn't every producer do so?), or whether these effects, while causal, reflect heterogeneous treatment effects, with firms most apt to benefit choosing to export. De Loecker (2007) finds that Slovenian firms that begin exporting during the post-transition period saw productivity growth after entering foreign markets. Interestingly, firms exporting to higher-income regions saw greater productivity growth. Apparently the export market—not just the exporter itself—matters. This raises interesting selection issues about which markets firms choose to export to, even conditional on the decision to export in the first place.

4.3. Deregulation or Proper Regulation

Poorly regulated markets can create perverse incentives that reduce productivity. Deregulating or reformatting to smarter forms of regulation can reverse this.

Benjamin Bridgman, Shi Qi, and Schmitz (2009) show how regulations in place for decades in the U.S. sugar market destroyed incentives to raise productivity. The U.S. Sugar Act, passed in 1934 as part of the Depression-era restructuring of agricultural law, funded a subsidy to sugar beet farmers with a tax on downstream sugar refining. Refiners were compensated for this tax by quota protection from imports and government-imposed limits on domestic competition (antitrust law was often thrown to the wind in the construction of New Deal programs). This transfer scheme led to the standard quantity distortions, but it also distorted incentives for efficient production. Farmers received a flat payment per ton of sugar contained in their beets, so their optimal response was to simply grow the largest beets possible. The problem is that refining larger beets into sugar is less efficient. As beets grow larger, their sugar-to-pulp ratio falls, requiring more time and energy to extract a given amount of sugar from them. At the same time, given the restraints on competition in the refined sugar market, refiners had little incentive to improve sugar extraction on the margin. The combined result of these incentives is readily apparent in the data. When the Sugar Act was passed, a ton of beets yielded an average of 310 pounds of refined sugar, a figure that had been steadily rising from 215 pounds per ton in 1900. But this trend suddenly reversed after 1934. Yields dropped to 280 pounds per ton by 1950 and 240 pounds by 1974, the year the Act was repealed. Not surprisingly, yields began to climb again immediately after repeal, to about 295 pounds per ton by 2004. It is a sad testimony to the

Act's productivity distortions that yields 70 years after the act were still lower than when it was passed.

Christopher R. Knittel (2002) and Kira Fabrizio, Nancy Rose, and Wolfram (2007) examine how power plant operations react to changes in the regulatory structure they operate under. Both studies involve moving plants away from a traditional cost-plus regulated monopoly structure into alternative forms. Knittel (2002) studies the implementation of "incentive regulation" programs, where regulators explicitly tie operators' earnings to the achievement of particular operating efficiencies. Fabrizio et al. look the effect of electricity market reforms that occurred in many regions in the U.S. during the 1990s. Both studies find that plants experienced efficiency gains after the shift in the regulatory environment. Fabrizio et al. also show that, in line with what one would expect, the productivity gains were largest among investor-owned utilities and smallest in municipally-operated utilities.

Beyond these case studies, recent work has also taken a broader look at how product market regulations impact productivity at the micro level. For example, Greenstone, List, and Syverson (2009) show how environmental regulations (the U.S. Clean Air Act Amendments specifically) reduce manufacturing plants' productivity levels. Nicoletti and Scarpetta (2005) and Jens Arnold, Nicoletti, and Scarpetta (2008) discuss the productivity effects of product-market regulations in OECD economies. A related yet distinct relation between legal structure and productivity is how privatization affects formerly state owned firms. J. David Brown, John S. Earle, and Almos Telegdy's (2006) study of formerly state owned enterprises in several Eastern European countries is one of the more comprehensive of such studies. They document broad-based productivity growth in plants after privatization, but they also find considerable variation in the size of the impacts across countries, with more than 15 percent average TFP growth in Romania, but a slightly negative impact in Russia.

4.4. Flexible Input Markets

I discussed above how competition increases productivity. If one thinks of competition as flexibility in product markets—in more competitive markets, it's easier for consumers to shift their purchases from one producer to another—it is logical to suppose that flexible *input* markets might also raise productivity levels.

Indeed, there are almost surely complementarities between product market and input market flexibility. If consumers want to reallocate their purchases across producers, firms that experience growth in demand for their products will need to hire additional inputs to meet that demand. The more easily inputs can move toward these firms, which will typically be higher-productivity businesses due to the forces described above, the faster and more smoothly the reallocation mechanism works. In the context of the model in Section 2, flexible input markets reduce the concavity of the revenue function, making producer size more responsive to productivity differences. This section discusses recent research tying factor market flexibility to productivity.

The institutional features of input markets, such as the roles of unions and the structure of the financial sector, have an ambiguous theoretical impact on flexibility. If institutions improve match efficiency, solve asymmetric information problems, or otherwise serve efficiency-enhancing roles, they make input markets more flexible. If they facilitate rent-seeking behavior on the other hand, they impede flexibility. In the end, the impact of any particular institution is an empirical question—one which several of the studies in this section investigate.

Maksimovic and Phillips (2001) investigate the market for U.S. manufacturing plants themselves, as productive assets. They measure how a plant's productivity changes when it is sold by one firm to another. They find that, on average, a plant's productivity rises after the sale. That is reassuring: the market tends to allocate inputs in an efficient way, instead of as a response to ambitions of empire-building managers or other inefficient motives. Another of their findings that is consistent with this efficiency-enhancing role is that the plants that are sold tend to come from the selling firm's less productive business lines. In essence, the sellers are moving away from activities at which they are less proficient.

Petrin and Sivadasan (2010) use a novel approach to look at the productivity effects of labor market flexibility. They measure the difference between Chilean plants' marginal products of labor (as derived from industry-level production functions they estimate) and their average wages. Such gaps can be caused by any one of a number of market distortions, like market power, taxes, or the firing costs that are the object of the study. Allocative efficiency is achieved, at least in the cross section, when this gap is equated across plants. (Though of course overall inefficiencies still exist unless these gaps are all zero.) Efficiency increases if labor inputs are moved from low- to high-gap plants, because the net change in marginal product caused by the input shift outstrips the change in wage costs. Petrin and Sivadasan find that a particular legislative change that raised firing costs was associated with an increase in the mean gap, suggesting the legislation reduced allocative efficiency.

Several recent papers have taken these ideas and asked whether, more broadly speaking, economies efficiently allocate inputs across heterogeneous production units. Hsieh and Klenow (2009) use the measured TFP dispersion across Chinese and Indian firms to infer the size of producer-level distortions that jointly depress aggregate productivity in those economies. Their methodology is conceptually similar to Petrin and Sivadasan's gap approach. Their model indicates that in the absence of distortions, plants' revenue-based TFP levels (TFP measured using revenues as an output measure rather than quantities) should be equal. This implies that observed deviations from this equality reflect the presence of distortions. (Note, however, that quantity-based TFP values are not equated even if there are no distortions.) Essentially, their framework implies that plants with relatively high (low) revenue TFP levels are too small (large) relative to an allocatively efficient benchmark.[17] After measuring these implied plant-level distortions, they compare their distribution to the analogous distribution measured in U.S. microdata. (This is used as the comparison rather than the first-best/zero-distortion outcomes because it is a more realistic control group. The U.S. data contain, and hence can be used to control for, gaps that reflect adjustment costs and measurement error that may be immutable to policy action.) Hsieh and Klenow find that Chinese aggregate TFP could increase by 30-50 percent and Indian TFP by 40-60 percent by achieving the U.S. level of allocative efficiency with their existing resources.

Bartelsman, Haltiwanger, and Scarpetta (2009) look at the success of allocation across several countries. Rather than using a gap-type methodology like Hsieh and Klenow, they measure efficiency using the correlation between a plant's share of industry output and its productivity level. The logic of this metric is straightforward and similar to that in the model in Section 2 and what was discussed at the beginning of the competition section. Well functioning markets should reallocate output to more productive plants, leading to a positive correlation

[17] Their model's implication of equal revenue TFP across plants stems from the standard efficiency condition that inputs' marginal revenue products are equated across all uses, and the fact that marginal products are proportional to average products for a Cobb-Douglas production function without fixed costs. Since TFP is an average product measure, equal marginal revenue products implies equal average revenue products and therefore equal revenue TFP. Non-Cobb-Douglas technologies and/or fixed costs can also support persistent revenue TFP differences aside from any distortions.

between output share and productivity. An additional advantage of the metric is that it is easy to compute. Its limitation is that it is an accounting decomposition, and as such is not directly tied to welfare theory the way gap-type measures are. However, Bartelsman et al. show in a simple model how various types of producer-level distortions do in fact lead to reductions in the output-productivity correlation within an industry.

5. Big Questions

That is a brief summary of what we know about the causes of productivity differences at the micro level, and why we would want to know these causes. I want to emphasize that while the discussion draws out major themes of that body of knowledge, it really only just scratches the surface of the literature.

I think a fair reading of the discussion above would say that we have learned a lot about productivity since the Bartelsman and Doms (2000) survey. At the same time, it is hardly time to declare victory and go home. Many pressing issues and open questions remain. In this section, I will briefly lay out what I see to be the major questions about productivity that the research agenda should address.[18]

What is the importance of demand? Productivity is typically thought of as a supply-side concept. As discussed in Section 2, it is the component of the production function unrelated to observable labor, capital, and intermediate inputs. But productivity as actually measured in producer microdata generally reflects more than just supply-side forces. Because producer-specific prices are unobserved in most business-level microdata, output is typically measured by revenue divided by an industry-level deflator. This means that within-industry price differences are embodied in output and productivity measures. If prices reflect in part idiosyncratic demand shifts or market power variation across producers—a distinct likelihood in many industries—then high "productivity" businesses may not be particularly technologically efficient. Much of the literature described above therefore documents the joint influence of productivity *and* demand factors that show up in within-industry price variation.

A new strand of research has begun to extend the productivity literature to explicitly account for such idiosyncratic demand effects as well. These new frameworks—see

[18] Conversations with John Haltiwanger were very helpful in writing this section.

Sanghamitra Das, Roberts, and James R. Tybout (2007); Eslava et al. (2008); Foster, Haltiwanger, and Syverson (2008, 2009), and De Loecker (2009), for example—allow an additional and realistic richness in the market forces that determine producers' fates. The work to this point indicates that demand factors are indeed important. They exert a considerable influence on businesses' growth and survival. And while many of the basic results above that have been checked after adjusting for the supply-demand dichotomy have been robust, the results do suggest some reinterpretations of productivity effects as inferred from standard measures.

The scope of issues that this new line of research has addressed is still small, however. Demand could play an important role in many more settings that have been hidden to this point due to measurement issues. This is likely to be especially true when moving to sectors without well defined outputs (what exactly does Google produce, for example, and how should it be measured?). Unwinding this knot is a top priority.

What is the role of (or hope for) government policies that encourage productivity growth? Clearly, many of the productivity drivers discussed above can be influenced by government policies. This is especially true of the "external" drivers in the previous section—the elements of the market environment that can induce business to take actions to raise their productivity or that affect the Darwinian selection process that whittles out inefficient producers.

Several policy-related questions are prime targets for research. There have been many policy reforms (particularly in trade policy and market regulation design) that had plausibly productivity-enhancing effects. Many studies have evaluated specific reforms in isolation, taking the policy change as given. But a policy change, even one that moves in the right direction, may not necessarily be optimal. Alternative reforms, either in size or approach, might be more cost effective. Research has typically compared the effects of policy reforms to a null of no reform, but perhaps an equally important comparison is among possible reform alternatives. What type of reform is most effective for a given type of market or friction? What is the optimal size and timing of policy changes? These are the next set of questions the literature should chase in this area.

A related issue is why reforms, even if they are welfare enhancing in their productivity effects, don't always happen. There could be economic reasons for this. Established interests could be earning rents in the unreformed environment. They may be able to stave off reform, especially if its benefits are diffuse while its losses are concentrated. Characterizing the nature of these barriers to aggregate productivity gains—who wins, who loses, and by how much—could be fruitful.

Which productivity drivers matter most? The research described above has framed which factors might explain variation in productivity levels. The relative quantitative importance of each, however, is still unclear. Summarized succinctly, if we could easily measure these factors and add them to the production function, which would have the largest R^2?

Of course, it's quite likely that the quantitative impact of factors varies across industries or markets. A concomitant question, then, is which factors matter most in what sectors? Research that ties observable attributes of the industry's technology or demand structure to the quantitative importance of productivity-influencing factors would be an incredible advance in our ability to explain productivity growth.

What factors determine whether selection or within-producer growth is more important in a market/sector/industry? In many settings above, there was a prominent distinction between aggregate productivity growth coming from "within" (productivity growth at a given plant or firm) and "between" (reallocation-based selection across existing businesses or entry and exit) sources. Just as the literature still needs to characterize the relative quantitative contribution of various influences on producer-level efficiency, so too does it need to measure the relative importance of within and between components in explaining aggregate productivity growth.

We do know some patterns already. For example, aggregate productivity growth in the retail sector seems to be almost exclusively from reallocation, at least in the U.S. But of course the literature has covered nowhere near the full span of sectors and economies. More importantly, we do not yet have a good model of what sectoral features (again on either the supply or demand side) might determine the relative importance of each. Why is within-store productivity growth so small on average in retail, but not manufacturing, for example? Answering questions like this would go a long way to developing our understanding of how micro productivity differences drive the aggregate productivity movements.

What is the role of misallocation as a source of variation in emerging economies? Productivity differences explain much of the per capita income variation across countries. As seen above, recent research with producer microdata is building the case that a substantial portion of these productivity gaps arise from poor allocation of inputs across production units in developing countries.

In some ways, this is a hopeful finding: these countries could become substantially more productive (and raise their incomes) by simply rearranging the inputs they already have. Not everything hangs on some unattainable technologies that are out of reach.

On the other hand, the result also has discouraging elements. While research has identified misallocation as a source of the problem, it hasn't really pinned down exactly what distortions create gaps between the social marginal benefits and costs of inputs across production units. It is hard to implement policies that close these gaps and the variation between them (i.e., reallocate inputs more efficiently) without knowing the nature of the gaps in the first place.

That said, there has been some early progress on this front. Witness the efforts to tie misallocation to various labor market policies. Much remains to be done, however, and this is an important area for further effort.

What is the importance of higher variance in productivity outcomes? Some of the work above, particularly that focusing on the role of IT capital, suggests that the variance of productivity outcomes might be increasing at a very broad level. This has several implications. First, the operation of a business is a call option: poor outcomes are truncated because of the possibility of exit. The value of this option increases with a mean-preserving spread in outcomes. As such, higher variance should lead to more firms taking bets on potential productivity-increasing activities like IT investment. There is some evidence that this is happening, but the literature has yet to show this definitively. Second, if there is an upward trend in productivity dispersion, will the forces of selection stem this spread? If so, when? Will a shakeout be strong enough to drive dispersion back to its previous level? Third, is this increase in variance something specific about IT capital, or is it a broader feature of general purpose technologies? Historical evidence would be very informative here. For example, did the diffusion of the electric motor in the early 20th Century also increase in the variance in productivity outcomes across businesses? Or even when a particular industry experiences a revolution in its standard technology, does this lead to temporary increases in productivity dispersion followed by a shakeout?

Can we predict innovation based on market conditions? Here I speak of innovation broadly—product and process innovation, measured or unmeasured by formal R&D numbers. This question is in some ways a corollary to the one above about quantifying and predicting the split between within-producer and between-producer productivity growth. Within-productivity growth is in many cases not simply the passive accumulation of efficiency; it comes in part as a result of the active innovative efforts of producers. What market or technological factors determine how large innovative activity will be? Can we predict whether product or process innovation will dominate, based on market features?

The nature of intangible capital. Many of the primary drivers of productivity naturally create persistence in productivity levels at plants and firms. These include learning-by-doing; innovative efforts; and in many cases investment in higher quality managerial, labor, or capital inputs. An easy way to explain such persistence is to think of these productivity enhancements as resulting from producers' investments in intangible capital—know-how about their businesses that is embodied in the organization. This conceptual structure also highlights how productivity gains sourced in intangible capital can also be thought of, along with managerial and unobserved factor qualities, as arising from mismeasured inputs. If one really could measure intangible capital (which, alas, is inherently difficult given its nature), the productivity differences arising from such sources could be explained.

Understanding how such intangible capital stocks are built and sustained would shed light on many productivity-related issues for this reason. Such insights would also speak toward active literatures on the subject in macroeconomics and finance. How much uncertainty is inherent in intangible capital investment? What is the distribution of rates of return across producers, and what predicts them? Is intangible capital fully excludable, or are there spillovers to other firms? How well do R&D measures capture investment in intangibles? Are there other proxies that could augment such measures?

Management vs. managers. We know more about the role of management than before, but what about *managers*? Some good work on CEOs aside, we don't really know if good managerial practices matter enough to attain productivity gains, or whether they are complementary to the skills of those who implement them. If they are complements, what skills matter? Are they built by experience, tenure in the industry or on the job, education, or something else? Understanding these issues might also help to pin down the causal nature of management practices. If good management practices reflect in large part the fact that they are what good managers do, then the causal impact might be limited. On the other extreme, if managers don't seem to matter at all, then it is quite likely that managerial practices have a strong causal impact on productivity.

A plea for data. Data availability is not a research question, but it is crucial for answering the questions posed above. Virtually everything discussed in this survey we now know because detailed data on production practices was available. But many of these data sets were originally collected by statistical agencies for the purpose of constructing aggregates. Their ability to offer insights into what happens at the micro level was in many ways a happy externality. Now that we know the value of the knowledge that such information can generate, economists should push for more directed efforts to measure business-level production practices. This could include, for example, more data on managers and management practices, business-level prices, input quality measures, proxies for intangible capital, non-R&D innovation spending, and so on. Obviously, collecting such data is costly, and this sort of push will involve tradeoffs for statistical agencies or a willingness of researchers to pay private companies for the collection efforts. Nevertheless, it seems clear that there is much to be gained in exchange for those costs.

6. Conclusion

The research into the productivity differences across businesses has come a long way since Bartelsman and Doms (2000) surveyed the literature a decade ago. We know more about what causes the measured differences in productivity, and how factors both internal and external to the plant or firm shape the distribution. These insights have been applied to research questions in numerous fields.

That said, there is still plenty to be learned. Fortunately, I see no sign that the rate at which researchers accumulate knowledge in this area is slowing. I am excited to see what the next several years bring in this research agenda, as the content of 2020's survey unfolds.

References

Abowd, John, John Haltiwanger, Ron Jarmin, Julia Lane, Paul Lengermann, Kristin McCue, Kevin McKinney, and Kristin Sandusky. 2005. "The Relationship Between Human Capital, Productivity and Market Value: Building Up From Microeconomic Evidence." In *Measuring Capital in the New Economy* (eds. Carol Corrado, John Haltiwanger and Daniel Sichel), Chicago: University of Chicago Press.

Ábrahám, Árpád and T. Kirk White. 2006. "The Dynamics of Plant-Level Productivity in U.S. Manufacturing." Center for Economic Studies Working Paper 06-20.

Abowd, John M., Francis Kramarz, David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251-333.

Acemoglu, Daron, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti. 2007. "Technology, Information, and the Decentralization of the Firm." *Quarterly Journal of Economics*, 122(4): 1759-99.

Acemoglu, Daron and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics*, 119(3): 1049-90.

Ackerberg, Daniel, C. Lanier Benkard, Steven Berry, and Ariel Pakes. 2007. "Econometric Tools for Analyzing Market Outcomes." in *Handbook of Econometrics: Vol. 6A*, eds. J. Heckman and E. Leamer, North Holland.

Amiti, Mary and Jozef Konings. "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia." *American Economic Review*, 97(5): 1611-38.

Arnold, Jens, Giuseppe Nicoletti, and Stefano Scarpetta. 2008. "Regulation, Allocative Efficiency and Productivity in OECD Countries: Industry and Firm-level Evidence." OECD Economics Department Working Paper 616.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Inventions." In *The Rate and Direction of Inventive Activity*, R. Nelson, ed. Princeton, N.J.: Princeton University Press.

Asplund, Marcus and Volker Nocke. 2006. "Firm Turnover in Imperfectly Competitive Markets." *Review of Economic Studies*, *73*(2): 295-327.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. Forthcoming. "R&D Investment, Exporting, and Productivity Dynamics." *American Economic Review*.

Balasubramanian, Natarajan and Jagadeesh Sivadasan. Forthcoming. "What Happens When Firms Patent? New Evidence from U.S. Economic Census Data." *Review of Economics and Statistics*.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2007. "Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment." *Quarterly Journal of Economics*, 122(2): 729-73.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2009. "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." *Econometrica*, 77(4): 1047-94.

Bartel, Ann, Casey Ichniowski and Kathryn Shaw. 2007. "How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills." *Quarterly Journal of Economics, 122*(4): 1721-1758.

Bartelsman, Eric and Mark Doms. 2000. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature*, *38*(3): 569-594.

Bartelsman, Eric, Pieter Gautier, and Joris de Wind. 2009. "Employment Protection, Technology Choice, and Worker Allocation." Working Paper.

Bartelsman, Eric, John Haltiwanger and Stefano Scarpetta. 2009. "Cross Country Differences in Productivity: The Role of Allocation and Selection." NBER Working Paper 15490.

Bartelsman, Eric J., Jonathan Haskel, and Ralf Martin. 2008. "Distance to Which Frontier? Evidence on Productivity Convergence from International Firm-level Data." CEPR Discussion Paper 7032.

Basu, Susanto, Luigi Pascali, Fabio Schiantarelli, and Luis Serven, 2009. "Productivity, Welfare and Reallocation: Theory and Firm-Level Evidence." NBER Working Paper 15579.

Benkard, C. Lanier. 2000. "Learning and Forgetting: The Dynamics of Aircraft Production." *American Economic Review*, 90(4): 1034-54.

Bernard, Andrew, J. Bradford Jensen and Peter Schott. 2006. "Trade Costs, Firms and Productivity." *Journal of Monetary Economics*, 53(5): 917-37.

Bernard, Andrew, Stephen Redding, and Peter Schott. 2010. "Multiple-Product Firms and Product Switching." *American Economic Review*, 100(1): 70-97.

Bertrand, Marianne and Antoinette Schoar. 2003. "Managing with Style: The Effect of Managers on Firm Policies," *Quarterly Journal of Economics*, *118*(4): 1169–1208.

Bloom, Nicholas, Mirko Draca, and John Van Reenen. 2009. "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, Diffusion and Productivity." LSE/Stanford Working Paper.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2010. "Does Management Matter? Evidence from India." Stanford Working Paper.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun and John Van Reenen. 2010. "Does Management Matter? New Empirics and Old Theories." LSE/Stanford Working Paper.

Bloom, Nicholas, Rafaella Sadun, and John Van Reenen. 2009a. "Americans Do I.T. Better: US Multinationals and the Productivity Miracle." LSE Working Paper.

Bloom, Nicholas, Rafaella Sadun, and John Van Reenen. 2009b. "The Organization of Firms Across Countries." NBER Working Paper 15129.

Bloom, Nicholas, Mark Schankerman and John Van Reenen. 2007. "Identifying Technology Spillovers and Product Market Rivalry." NBER Working Paper 13060.

Bloom, Nicholas and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *Quarterly Journal of Economics*, 122(4): 1351-1408.

Bloom, Nicholas and John Van Reenen. 2010. "Why Do Management Practices Differ across Firms and Countries?" *Journal of Economic Perspectives*, 24(1): 203-24.

Blundell, Richard and Stephen R. Bond. 2000. "GMM Estimation with Persistent Panel Data: An Application to Production Functions." *Econometric Reviews*, 19(3): 321-40.

Boning, Brent, Casey Ichniowski, and Kathryn Shaw. 2007. "Opportunity Counts: Teams and the Effectiveness of Production Incentives." *Journal of Labor Economics*, *25*(4): 613-50.

Bridgman, Benjamin, Shi Qi, and James A. Schmitz, Jr. 2009. "The Economic Performance of Cartels: Evidence from the New Deal U.S. Sugar Manufacturing Cartel, 1934-74." Federal Reserve Bank of Minneapolis, Research Department Staff Report 437.

Brown, J. David, John S. Earle, and Almos Telegdy. 2006. "The Productivity Effects of Privatization: Longitudinal Estimates from Hungary, Romania, Russia, and Ukraine." *Journal of Political Economy*, 114(1): 61-99.

Brynjolfsson, Erik, Andrew McAfee, Michael Sorell, and Feng Zhu. 2007. "Scale without Mass." Harvard Business School Working Paper 07-016.

Bushnell, James B. and Catherine Wolfram. 2009. "The Guy at the Controls: Labor Quality and Power Plant Efficiency." in *International Differences in the Business Practices and the Productivity of Firms*, eds. Richard Freeman and Kathryn Shaw, Chicago: University of Chicago Press.

Campbell, Jeffrey R. and Jonas D. M. Fisher. 2004. "Idiosyncratic Risk and Aggregate Employment Dynamics." *Review of Economic Dynamics*, 7(2): 331-53.

Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert. 1982. “The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity.” *Econometrica, 50*(6): 1393-1414.

Collard-Wexler, Allan. 2008 “Demand Fluctuations in the Ready-Mix Concrete Industry.” Working Paper.

Cooper, William W., Lawrence M. Seiford, and Kaoru Tone. 2006. *Introduction to Data Envelopment Analysis and Its Uses*. New York: Springer.

Crespi, Gustavo, Chiara Criscuolo, Jonathan E. Haskel, and Matthew Slaughter. “Productivity Growth, Knowledge Flows, and Spillovers.” NBER Working Paper No. 13959.

Cummins, Jason G. and Giovanni L. Violante. 2002. “Investment-Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences.” *Review of Economic Dynamics, 5*(2): 243-84.

Das, Sanghamitra, Mark J. Roberts, and James R. Tybout. “Market Entry Costs, Producer Heterogeneity, and Export Dynamics.” *Econometrica*, *75*(3), 2007, 837-73.

De Loecker, Jan. 2007. “Do Exports Generate Higher Productivity? Evidence from Slovenia.” *Journal of International Economics*, *73*(1), 69-98.

De Loecker, Jan. 2009. “Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity.” Working Paper.

Disney, Richard, Jonathan Haskel, and Ylva Heden. 2003a. “Entry, Exit and Establishment Survival in UK Manufacturing.” *Journal of Industrial Economics*, *51*(1): 91-112.

Disney, Richard, Jonathan Haskel, and Ylva Heden. 2003b. “Restructuring and Productivity Growth in UK Manufacturing.” *Economic Journal*, *113*(489): 666-94.

Doraszelski, Ulrich and Jordi Jaumandreu. 2009. “R&D and Productivity: Estimating Endogenous Productivity.” Working paper.

Eaton, Jonathan and Samuel Kortum. 2002. “Technology, Geography, and Trade.” *Econometrica, 70*(5): 1741-79.

Ericson, Richard and Ariel Pakes. 1995. “Markov-Perfect Industry Dynamics: A Framework for Empirical Work.” *Review of Economic Studies, 62*(1): 53-82.

Eslava, Marcela, John C. Haltiwanger, Adriana Kugler and Maurice Kugler. 2004. “The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia.” *Journal of Development Economics, 75*(2): 333-72.

Eslava, Marcela, John C. Haltiwanger, Adriana Kugler and Maurice Kugler. 2008. “Plant

Survival, Market Fundamentals and Trade Liberalization." Working Paper.

Fabrizio, Kira, Nancy Rose and Catherine Wolfram. 2007. "Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on U.S. Electric Generation Efficiency." *American Economic Review*, *97*(4), 1250-1277.

Faggio, Giulia, Kjell G. Salvanes, and John Van Reenen. 2009. "The Evolution of Inequality in Productivity and Wages: Panel Data Evidence." LSE Working Paper.

Fernandes, Ana M. 2007. "Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries." *Journal of International Economics*, 71(1): 52-71.

Forbes, Silke and Mara Lederman. 2009. "Does Vertical Integration Affect Firm Performance? Evidence from the Airline Industry." Working Paper.

Foster, Lucia, John Haltiwanger and C.J. Krizan. 2001. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." *NBER Studies in Income and Wealth, vol. 63: New Developments in Productivity Analysis*. Chicago and London: University of Chicago Press, 303-63.

Foster, Lucia, John C. Haltiwanger and C.J. Krizan. 2006. "Market Selection, Restructuring and Reallocation in the Retail Trade Sector in the 1990s." *Review of Economics and Statistics*, 88(4): 748-58.

Foster, Lucia, John C. Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, *98*(1): 394-425.

Foster, Lucia, John C. Haltiwanger, and Chad Syverson. 2009. "The Slow Growth of New Plants: Learning about Demand?" Working Paper.

Fox, Jeremy and Valerie Smeets. 2009. "Does Input Quality Drive Measured Differences in Firm Productivity?" Working Paper.

Galindo-Rueda, Fernando and Jonathan Haskel. 2005. "Skills, Workforce Characteristics and Firm-Level Productivity in England." Report prepared for the Department of Trade and Industry, Department for Education and Skills, Office for National Statistics.

Garicano, Luis and Paul Heaton. 2007. "Information Technology, Organization, and Productivity in the Public Sector: Evidence from Police Departments." LSE Centre for Economic Performance Working Paper.

Gersbach, Hans and Armin Schmutzler. 2003. "Endogenous Technological Spillovers: Causes and Consequences." *Journal of Economics and Management Strategy*, 12(2): 179-205.

Greenstone, Michael, John A. List, and Chad Syverson. 2009. "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing." Working Paper.

Griffith, Rachel, Rupert Harrison and John Van Reenen. 2007. "How Special is the Special Relationship? Using the Impact of U.S. R&D Spillovers on U.K. Firms as a Test of Technology Sourcing." *American Economic Review*, 96(5): 1859-75.

Griliches, Zvi and Jacques Mairesse. 1998. "Production Functions: The Search for Identification." *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*. Cambridge, New York and Melbourne: Cambridge University Press: 169-203.

Haltiwanger, John, Stefano Scarpetta, and Helena Schweiger. 2008. "Assessing Job Flows Across Countries: The Role of Industry, Firm Size and Regulations." NBER Working Paper 13920.

Hamilton, Barton H., Jack A. Nickerson and Hideo Owan. 2003. "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation." *Journal of Policital Economy*, *111*(3): 465-497.

Holmes, Thomas J., David K. Levine, and James A Schmitz Jr. 2008. "Monopoly and the Incentive to Innovate When Adoption Involves Switchover Disruptions." NBER Working Paper 13864.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, *60*(5): 1127-50.

Hortacsu, Ali and Chad Syverson. 2007. "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." *Journal of Political Economy*, *115*(2): 250-301.

Hortaçsu, Ali and Chad Syverson. 2009. "Why Do Firms Own Production Chains?" Working Paper.

Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403–48.

Hubbard, Thomas N. 2003. "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking." *American Economic Review*, 93(4): 1328-53.

Hulten, Charles R. 2009. "Growth Accounting." NBER Working Paper 15341.

Ichniowski, Casey and Kathryn Shaw. 2003. "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices." *Journal of Economic Perspectives*, *17*(1): 155-78.

Ichniowski, Casey, Kathryn Shaw and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *American Economic Review*, *87*(3): 291-313.

Ilmakunnas, Pekka, Mika Maliranta, and Jari Vainiomäki. 2004. “The Roles of Employer and Employee Characteristics for Plant Productivity.” *Journal of Productivity Analysis*, 21(3): 249-276.

Jorgenson, Dale W., Mun S. Ho and Kevin J. Stiroh. 2005. *Information Technology and the American Growth Resurgence*. Cambridge: MIT Press.

Jorgenson, Dale W., Mun S. Ho and Kevin J. Stiroh. 2008. “A Retrospective Look at the U.S. Productivity Growth Resurgence.” *Journal of Economic Perspectives*, *22*(1): 3-24.

Jovanovic, Boyan. 1982. “Selection and the Evolution of Industry.” *Econometrica*, *50*(3): 649-70.

Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen. 2009. “Which CEO Characteristics and Abilities Matter?” University of Chicago Working Paper.

Keller, Wolfgang and Stephen R. Yeaple. 2009. “Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States.” *Review of Economics and Statistics*, 91(4): 821-31.

Kellogg, Ryan. 2009. “Learning by Drilling: Inter-Firm Learning and Relationship Persistence in the Texas Oilpatch.” NBER Working Paper 15060.

Klette, Tor Jakob and Samuel Kortum. 2004. “Innovating Firms and Aggregate Innovation.” *Journal of Political Economy*, *112*(5): 986–1018.

Knittel, Chirstopher R. 2002. “Alternative Regulatory Methods and Firm Efficiency: Stochastic Frontier Evidence the US Electricity Industry.” *Review of Economics and Statistics*, 84(3): 530-540.

Lazear, Edward P. 2000. “Performance Pay and Productivity.” *American Economic Review*, *90*(5): 1346-61.

Lentz, Rasmus and Dale T. Mortensen. 2009. “An Empirical Model of Growth through Product Innovation.” *Econometrica*, *76*(6): 1317–73.

Levinsohn, James and Amil Petrin. 2003. “Estimating Production Functions Using Inputs to Control for Unobservables.” *Review of Economic Studies*, *70*(2): 317-41.

Levitt, Steven, John List, and Chad Syverson. 2010. “How Does Learning by Doing Happen?” Working Paper.

Maksimovic, Vojislav and Gordon Phillips. 2002. “Do Conglomerate Firms Allocate Resources Inefficiently across Industries? Theory and Evidence.” *Journal of Finance*, 57(2), 721-767.

Maksimovic, Vojislav and Gordon Phillips. 2001. “The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?” *Journal of Finance*, 56(6), 2019-2065

Malmendier, Ulrike and Geoffrey Tate. 2009. “Superstar CEOs.” *Quarterly Journal of Economics,* 124(4): 1593–1638.

Marschak, Jacob and William H. Andrews, Jr. 1944. “Random Simultaneous Equations and the Theory of Production.” *Econometrica*, *12*(3/4): 143-205.

Mas, Alexandre. 2008. “Labour Unrest and the Quality of Production: Evidence from the Construction Equipment Resale Market.” *Review of Economic Studies*, 75(1): 229-58.

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, 71(6): 1695-1725.

Moretti, Enrico. 2004. “Workers’ Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions.” *American Economic Review*, *94*(3): 656-690.

Muendler, Marc A. 2004. “Trade, Technology, and Productivity: A Study of Brazilian Manufacturers, 1986-1998.” CESifo Working Paper, 1148.

Nicoletti, Giuseppe and Stefano Scarpetta. 2005. “Regulation and Economic Performance: Product Market Reforms and Productivity in the OECD.” OECD Economics Department Working Paper No. 460.

Oliner, Stephen D., Daniel Sichel, and Kevin Stiroh. 2007. “Explaining a Productive Decade.” *Brookings Papers on Economic Activity*, 2007 Issue 1: 81-137.

Olley, G. Steven and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica*, 64(6), 1263-97.

Organization for Economic Cooperation and Development. 2001. “Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth.” *OECD Manual*, Paris: OECD Publications.

Pavcnik, Nina. 2002. “Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants.” *Review of Economic Studies*, 69(1): 245-76.

Petrin, Amil and James Levinsohn. 2008. “Measuring Aggregate Productivity Growth Using Plant-Level Data.” Working Paper no. 11887.

Petrin, Amil and Jagadeesh Sivadasan. 2010. “Job Security Does Affect Economic Efficiency: Theory, a New Statistic, and Evidence from Chile.” Working paper.

Rose-Ackerman, Susan. 1991. “Risk Taking and Ruin: Bankruptcy and Investment Choice.” *Journal of Legal Studies*, 20(2): 277-310.

Sakellaris, Plutarchos and Daniel J. Wilson. 2004. “Quantifying Embodied Technological Change.” *Review of Economic Dynamics*, *7*(1): 1-26.

Schmitz, James A., Jr. 2005. “What Determines Productivity? Lessons from the Dramatic Recovery of the U.S. and Canadian Iron Ore Industries following Their Early 1980s Crisis.” *Journal of Political Economy*, 113(3), 582-625.

Schoar, Antoinette. 2002. “The Effect of Diversification on Firm Productivity.” *Journal of Finance*, 62(6): 2379-2403.

Shearer, Bruce. 2004. “Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment.” *Review of Economic Studies*, 71(2), 513-34.

Syverson, Chad. 2004a. “Product Substitutability and Productivity Dispersion.” *Review of Economics and Statistics*, 86(2): 534-550.

Syverson, Chad. 2004b. “Market Structure and Productivity: A Concrete Example.” *Journal of Political Economy*, 112(6): 1181-1222.

Syverson, Chad. 2007. “Prices, Spatial Competition, and Heterogeneous Producers: An Empirical Test.” *Journal of Industrial Economics*, 55(2): 197-222.

Thornton, Rebecca Achee and Peter Thompson. 2001. “Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding.” *American Economic Review*, 91(5): 1350-68.

UK Office of Fair Trading. 2007. “Productivity and Competition: An OFT Perspective on the Productivity Debate.”

van Ark, Bart, Mary O’Mahony, and Marcel P. Timmer. 2008. “The Productivity Gap between Europe and the United States: Trends and Causes.” Journal of Economic Perspectives, *22*(1): 25-44.

Van Biesebroeck, Johannes. 2003. “Productivity Dynamics with Technology Choice: An Application to Automobile Assembly.” *Review of Economic Studies*, 70(1): 167-98.

Van Biesebroeck, Johannes. 2005. “Exporting Raises Productivity in sub-Saharan African Manufacturing Firms.” *Journal of International Economics*, 67(2): 373-91.

Van Biesebroeck, Johannes. 2008. “The Sensitivity of Productivity Estimates: Revisiting Three Important Productivity Debates.” *Journal of Business and Economic Statistics*, 26(3): 321-38.

Van Reenen, John. 1996. “The Creation and Capture of Rents: Wages and Innovation in a Panel of U.K. Companies.” *Quarterly Journal of Economics*, 111(1): 195-226.

Veracierto, Marcelo. 2008. “Firing Costs and Business Cycle Fluctuations.” *International Economic Review*, 49(1): 1-39.

Verhoogen, Eric A. 2008. “Trade, Quality Upgrading and Wage Inequality in the Mexican Manufacturing Sector.” *Quarterly Journal of Economics*, 123(2): 489-530.

Vives, Xavier. 2007. “Innovation and Competitive Pressure.” IESE Business School Working Paper no. 634.

Walker, Francis A. 1887. “The Source of Business Profits.” *Quarterly Journal of Economics*, *1*(3): 265-288.