On The Micro-Foundations of Productivity Growth

by

James Levinsohn
National Bureau of Economic Research
University of Michigan
and
Amil Petrin
University of Chicago, GSB
National Bureau of Economic Research

Current version: November 21, 2003

Abstract. We show that the traditional approach to aggregating plant-level productivity has no well-defined unit of measurement. We propose a simple measure that has a readily interpreted economic magnitude. We describe conditions that a decomposition of plant-level productivity growth must satisfy in order to identify rationalization effects separately from real productivity effects. We investigate the importance of the concerns we raise, showing that the seemingly innocuous choice of a decomposition can actually *reverse* the economic conclusions one draws. We provide new suggestions for exploring micro-foundations that are complementary to the usual decompositions, and which can be particularly useful when these decompositions fail. Finally, we use recent Chilean data spanning 10 years (1987-96) to illustrate our concerns.

Address. Levinsohn: Department of Economics and Ford School of Public Policy, University of Michigan, Ann Arbor, MI 48109; Internet: jamesL@umich.edu, WWW: http://www.econ.lsa.umich.edu/~jamesl, Petrin: Graduate School of Business, University of Chicago, 1101 E. 58th St., Chicago, IL 60637; Internet: amil.petrin@gsb.uchicago.edu, WWW:http://gsbwww.uchicago.edu/fac/amil.petrin

On The Micro-Foundations of Productivity Growth

James Levinsohn
National Bureau of Economic Research
University of Michigan
and
Amil Petrin
University of Chicago, GSB
National Bureau of Economic Research

1. Introduction

As recently as a decade ago, most estimates of industry-level productivity were obtained from industry-level data. With the increasing availability of plant-level data, industry-level productivity is now being constructed from plant-level measures. Recent plant-level productivity estimates have been obtained using U.S. data from the Longitudinal Research Database (LRD) of the U.S. Census, (even better) French data from the Declarations Annuelles des Salaires (DAS) collected by INSEE (Institut National de la Statistique et des Etudes Economiques), and several plant-level manufacturing censes from developing countries.[1]

With the move from industry-level data to plant-level data, it becomes possible to investigate the micro-foundations of productivity dynamics. Research on this issue has typically focused on decomposing the industry-level productivity changes into components that can be computed using plant-level productivity estimates. In particular, one can explore the relative roles of plant-level productivity increases and the reallocation of output from less productive plants to more productive plants. Both serve to increase standard measures of industry productivity, although policy recommendations can differ dramatically depending upon which kind of growth is prevalent in an industry. Examples of this work, all based on U.S. data, include Bailey et al. (1992), Foster et al. (2001), and Bernard and Jensen (2001).

A predecessor to this paper appeared as "When Industries Become More Productive, Do Firms?: Investigating Productivity Dynamics," and is available as NBER Working Paper 6893. Ivan Kandilov, Wendy Petropoulos and Jagadeesh Sivadasan provided splendid research assistance and many good ideas. We are grateful to the Russell Sage Foundation for support. Stata code for all computations is available from the authors by request.

[1] Examples of papers using the U.S. data include Bailey, Hulten, and Campbell (1992), Olley and Pakes (1996), Bernard and Jensen (1999), Bernard, Eaton, Jenson, and Kortum (2000), and Foster, Haltiwanger, and Krizan (2001). Abowd, Kramarz, and Margolis (1999) and Eaton, Kortum, and Kramarz (2001) use the French data while the several papers in Roberts and Tybout (1996) use the LDC data. These are but examples. A careful bibliography would include dozens of papers using plant-level data.

The questions being asked in the existing research are important, as understanding the micro-foundations of aggregate productivity dynamics matters.[2] When industry-level productivity increases and plants are becoming more productive, this is certainly welfare improving; there is more output from the same inputs. However, industry-level productivity can increase even if every plant becomes *less* productive. This counter-intuitive outcome can occur if there is reallocation of inputs from less productive to relatively more productive plants (whose productivities themselves are falling.) If every plant were to produce less output for the same inputs yet industry productivity were to increase, the news is, at best, mixed. There is a benefit to society from shifting inputs from less productive plants to more productive ones, but declining plant-level productivity is hardly good news for long-run economic growth. This paper is about trying to better understand just what underlies changes in industry-level growth.

We begin with the issue of aggregating plant-level productivity estimates to an industry-level measure. As described in Foster et al. (2001), the current methodology of choice is to average plant-level productivity residuals using either output shares or input shares as a plant's weight in the aggregation. We show that share-weighted approaches to aggregation lead to productivity measures that have no well-defined unit. This presents a problem when one wants to measure the value of real productivity changes and to compare productivity growth across time, industries, and countries. We recommend a simple alternative that does not suffer from this shortcoming. It uses the historical definition of productivity - more output holding inputs constant - and the empirical results illustrate how having meaningful units provides for results that are easier to interpret.

We then revisit the question of what underlies aggregate productivity growth (regardless of the units in which it is measured.) We expand on an issue originally raised in Foster et al. (2001), describing conditions that a decomposition of plant-level productivity growth must satisfy in order to identify rationalization effects separately from real productivity effects. We investigate the importance of these concerns, showing that the seemingly innocuous choice of a decomposition can actually *reverse* the economic conclusions one draws. We provide new suggestions for exploring micro-foundations that are complementary to the usual decompositions, and which can be particularly useful when these decompositions fail. Finally, we use recent Chilean data spanning 10 years

[2] Along with the voluminous recent empirical literature decomposing productivity growth into its components, there is a concurrent and closely related theoretical literature that models the micro-foundations of industry productivity growth. This literature (discussed below) does not provide uniform predictions about when the practitioner should expect to observe real productivity growth to dominate reallocation or vice versa. Indeed, the sheer variety of results from the theoretical literature speaks to the importance of empirically distinguishing the components of productivity growth.

(1987-96) to show, by example, that analyses of what underlies industry-level productivity changes often deserve a more skeptical viewing than that which they are usually accorded.

The remainder of this paper is organized as follows. In the next section, we describe a number of important policy questions that can only be answered using a productivity decomposition. In Section 3, we introduce a new metric for measuring industry-level productivity, and in Section 4 and 5 we describe some analytic frameworks for looking at what underlies industry-level productivity dynamics. In Section 6 we provide a short discussion of the data we employ. Section 7 describes the estimation and section 8 discusses the results.

2. Real Productivity vs. Reallocation: Why We Care

We term the notion that industry productivity increases because plants become more productive the "real productivity case." Plants might become more productive because they learn by doing or because they are exposed to new and better methods of production. Examples of the former range from the low-tech sewing machine operator at an apparel firm to the high-tech process of increasing yields on silicon chip production. An example of the latter is the set of manufacturing practices known as "lean production." U.S. manufacturers, initially auto producers but later others, adopted these methods after observing Japanese success.[3] In both the learning-by-doing and the learning-by-watching cases, firm productivity increases and with that comes increases in industry productivity. This is an uplifting explanation of the mechanism for productivity increases, as there are no obvious bounds to learning and ingenuity.

Industry productivity can also change because inputs are reallocated. In open markets, some firms thrive while others disappear. There are, within an industry, winners and losers. As firms that are especially well suited to an industry expand and misfits contract or exit, industry productivity increases. Conversely, industry productivity is hindered when firms are sheltered from the harsh realities of the marketplace.[4] This process of industry rationalization might be expected to lead to increased industry productivity, and we refer to this story as the rationalization case. Allowing for the rationalization case requires explicitly modeling firm heterogeneity, since a representative

[3] Biesebroeck (2001) examines this phenomenon and, like many of the previously cited studies, decomposes industry-level productivity changes into its micro-foundations.

[4] An example of this is an explanation given by the *Economist* for Japan's recent and long economic downturn. (See the June 20, 1998 issue containing the article "Japan's Economic Plight.") The *Economist*, echoing arguments made by many others, suggested that Japan's poor economic performance was, at least in part, because of a Japanese aversion to "outright failure" of firms. Rather, there exists a corporate culture more willing to tolerate low returns. In today's more global economy, this culture is more difficult to maintain and has contributed, the *Economist* claimed, to the recent poor performance of the Japanese economy while other industrial economies continued to boom.

firm framework (i.e. an aggregate production function) cannot adequately capture the evolutionary process in which some firms thrive while others lag.

Being able to separate the real productivity explanation for industry-level productivity changes from the rationalization explanation is important for several reasons. Perhaps most important from a policy standpoint is that the explanations have strikingly different implications for long run growth. In the real productivity case, industries become more productive because firms become more productive. This process is not bounded in any obvious way; a good idea can follow good idea after good idea. In the rationalization case (ignoring entry), an industry cannot become more productive than the single most productive incumbent firm. That is, the "frontier" is reached when all output is manufactured by the most productive firm, and productivity is constant thereafter.

From a policy perspective, prescriptions which might either promote or hinder productivity in the real productivity case often have different impacts when rationalization is more prevalent. Consider, for example, infant industry protection or subsidies to newly established firms. If these protected or subsidized plants become more productive over time, these policies may result in increased industry productivity and an internationally competitive industry. This outcome, though, requires increases in plant-level productivity– the real productivity case. If instead reallocation is what is driving changes in industry productivity, protecting or subsidizing new and less efficient plants only postpones the evolution toward a more productive industry.

Another policy-driven reason for identifying the relative roles of real productivity and rationalization concerns factor markets. The real productivity case and the rationalization case can have very different implications for these markets. The rationalization case entails substantial worker displacement, as jobs are moved from one firm to another in response to economic forces. By contrast, worker displacement may be a non-issue if all firms are becoming more productive at similar rates. With these differences in factor market implications come differences in the politics of productivity change.

Understanding whether growth is being driven by real productivity changes, reallocation, or an equal mix of the two is also helpful for evaluating the appropriateness of many theoretical models of firm dynamics. These models can be useful to the practitioner trying to understand what drives industry evolution, and they provide frameworks for researchers to conduct counter-factual policy analysis. Predictions for real productivity and reallocation are varied both within and across these models. In Jovanovic (1982), Hopenhayn and Rogerson (1993), Ericson and Pakes (1995), and Melitz (2001), there is entry, exit, and individual firms becoming more or less productive over time.

Out of equilibrium, there can be both reallocation and real growth in these models.[5] However, except for Ericson and Pakes (1995), in the equilibriums of these models, there is no role for reallocation in explaining productivity growth.[6]

As important as it may be to untangle the rationalization versus real productivity explanations, misinterpretations can be substantively important. In the "Results" section we illustrate some of the practices and pitfalls using recent Chilean manufacturing data.

3. Industry-level Productivity

Changes in the value of output per unit input is historically how productivity growth has been defined.[7] Total Factor Productivity (TFP), the standard measure of productivity, is most often defined as the difference between log output and predicted log output given log inputs for a Cobb-Douglas production technology.[8] With data indexed by plants i and time t, this technology is written as

$$y_{it} = \beta' x_{it} + \omega_{it} \quad (1),$$

where y_{it} is the log of Y_{it}, output or revenue measured in *levels*, and x_{it} is a vector of log inputs with coefficients β. TFP is given by ω_{it}, the multiplicative factor that scales $e^{\beta' x_{it}}$, an input index, to get final output.[9] Small changes in ω_{it} over time can be interpreted as percentage changes in i's output holding i's inputs constant.

When production data are aggregated beyond the plant-level - often all that is available to practitioners - an aggregate production technology must be well-defined to proceed. One implication of this technology is that any redistribution of inputs across plants results in the same aggregate output. Thus, the production function can be expressed as

$$y_t = \beta' x_t + \omega_t, \quad (2)$$

[5] In Melitz (2001), for example, reallocation plays the key role in the out-of-equilibrium transitions.

[6] The joint distribution of productivity and other state variables does not change over time, up to a common productivity growth term that is constant across firms.

[7] See Hulten (2001) for a careful and complete history of this index.

[8] The discussion here and below extends directly to any production technology.

[9] In some formulations, the disturbance term is comprised of two parts, and is written as $\omega_{it} + \eta_{it}$, where. e.g. η_{it} may reflect measurement error in the dependent variable like utilization that the researcher wishes to control for. Whether one measures productivity by ω_{it} alone or by the sum $\omega_{it} + \eta_{it}$ is a modeling decision and does not affect the discussion here.

with $y_t = ln(Y_t)$, $Y_t = \sum_i Y_{it}$, and similarly for x_t.[10] Here, holding inputs constant at the $t-1$ level, the change in the level of *aggregate* output from $t-1$ to t due to productivity growth is given as

$$\Delta Y_t = \Delta\omega_t * Y_{t-1}, \tag{3}$$

where $\Delta Y_t = Y_t - Y_{t-1}$ and $\Delta\omega_t$ is similarly defined. Historically, this is how aggregate productivity growth has been defined, with $\Delta\omega_t$ interpreted as the percentage change in industry output holding inputs constant.

When this aggregation condition does not hold, β must be estimated using plant-level data, that is, using (1) instead of (2).[11] The failure of aggregation raises a new methodological issue: how should industry-level productivity be defined in terms of the plant-level data?

Three rules of thumb have established themselves in the productivity literature using micro data. First, the appropriate units for the plant-level productivity term in the aggregation are the exponentiated log-levels, or

$$\tilde{\omega}_{it} = e^{\omega_{it}}.$$

Second, a plant's weight in the aggregation should reflect its economic size, so larger plants contribute more to the aggregated productivity number than do smaller plants. Finally, these plant weights should sum to one.[12] Together these guidelines yield the most commonly used measure of industry-level productivity (call it I_t):

$$I_t = \sum_{i=1}^{N} s_{it}\tilde{\omega}_{it}, \tag{4}$$

where s_{it} is plant i's share of either output or input in year t and $\tilde{\omega}_{it}$ is the estimate of plant-level productivity. Simply put, (4) is the average of (exponentiated) plant-level TFP residuals.

Although widely used, this aggregated quantity is hard to interpret because its unit of measurement is not well-defined. Consider trying to use a time series of these numbers to estimate

[10] Another implication is that marginal costs must be constant across all levels of output.

[11] We assume a consistent methodology exists for estimation with the plant-level data (see, for example, Levinsohn and Petrin (2003) if simultaneity is a problem).

[12] The traditional approach is to weight by the plant's share of industry output.Foster et al. (2001), however, note that output shares are not independent of productivity (i.e. in the production function, output is a function of inputs and a residual interpreted as productivity.) They argue for alternatively weighting observations by input shares. If one uses an input to weight plant productivity, industry rationalization will be interpreted as a shifting of *inputs* from less productive plants to more productive plants.

the change in the level of aggregate output due to productivity growth. Using the production technology from (1) to proceed, the level of a plant's output is given by

$$Y(X, \omega; \beta) = e^{\beta' x + \omega}.$$

Let the plant-level input index, be

$$z_{it} = e^{\beta' x_{it}}.$$

Then, holding each plant's input index constant at the initial period level $z_{i,t-1}$, a change in each plant's productivity from $\tilde{\omega}_{i,t-1}$ to $\tilde{\omega}_{it}$ leads to an aggregate output change of

$$\Delta Y_t = \sum_i z_{i,t-1} \Delta \tilde{\omega}_{it}. \tag{5}$$

Generally speaking, no function of the I_t series can reproduce ΔY_t because I_t uses weights different from $z_{i,t-1}$. Put another way, averaging the $\tilde{\omega}_{it}$ using share-weights does not aggregate in a manner that reflects the relationship between a plant's input level and its output level (from the production technology given in (1)). This means, for example, that the change in industry output holding industry inputs constant is *not* given by multiplying Y_{t-1} by ΔI_t.[13] This shortcoming of I_t also makes comparisons of productivity growth across industries difficult, as our results demonstrate later.

For the measurement of real productivity growth we propose the simple-to-compute alternative given in (5). By construction it satisfies the historical definition of productivity and has the units of output. More generally, for the decompositions given later, we suggest using (4) but with z_{it} in place of s_{it}, yielding a measure we call P_t:

$$P_t = \sum_{i=1}^{N} z_{it} \tilde{\omega}_{it}. \tag{6}$$

Every element in the sum entering P_t and (5) has as units the natural (original) unit given by the dependent variable Y_{it}, so both P_t and (5) also have these units (e.g. Dollars or Pesos). This facilitates measurement of the magnitude of economic growth over time and across industries. Finally, P_t has the feature that it aggregates to a predicted measure of industry output in levels, so changes in P_t reflect changes in the level of output over time that arise because real productivity changes, the level of industry inputs changes, and the allocation of inputs across firms changes.

[13] Usually, researchers normalize I_t to a base year and then measure percentage changes in the normalized index to summarize industry productivity growth. Normalization of an already uninterpretable index does not "fix" the interpretation problem. Rather, it remains unclear just how to interpret the economic magnitude of these percentage changes because the base level I_0 itself has no clearly defined units.

Having defined the traditional and an alternative units-based measure of industry productivity, we turn to the task of understanding the micro-foundations underlying the changes in industry-level growth.

4. Decomposing Productivity and Output Changes

A decomposition that is designed to identify the rationalization and real productivity roles in growth ideally should meet three conditions. First, both the real productivity and the rationalization cases are about *changes* in plant-level outcomes, so measurement needs to account for changes over time. Second, the measurement of the real productivity case should not be contaminated with realized changes in rationalization outcomes. Third, measurement of rationalization should not be contaminated with realized changes in plant-level productivity (i.e. real productivity outcomes.) A decomposition satisfying these conditions will, in many cases, be able to determine the relative roles of both growth mechanisms, and also have power to distinguish between the theory models described earlier.

The most popular decomposition is based on changes in I_t, the most popular measure of industry productivity. One can decompose $\Delta I_t = I_t - I_{t-1}$ according to the following:

$$\begin{aligned}\Delta I_t = &\sum_{i \in C} s_{i,t-1} \Delta \tilde{\omega}_{it} + \sum_{i \in C} \tilde{\omega}_{i,t-1} \Delta s_{it} + \sum_{i \in C} \Delta s_{it} \Delta \tilde{\omega}_{it} + \\ &\sum_{i \in B} s_{it} \tilde{\omega}_{it} - \sum_{i \in D} s_{i,t-1} \tilde{\omega}_{i,t-1}\end{aligned} \tag{7}$$

where C is the set of continuing plants, B the set of entrants, and D the set of exiters, and the difference operator, Δ, denotes the difference between year t and $t-1$. The decomposition has four components, in order: a real productivity growth term, and rationalization term, a covariance term, and a net entry term. Since ΔI_t has no clearly defined units, we propose a complementary decomposition to ΔI_t, that given by decomposing P_t, which has the units of the dependent variable and is given by

$$\begin{aligned}\Delta P_t = &\sum_{i \in C} z_{i,t-1} \Delta \tilde{\omega}_{it} + \sum_{i \in C} \tilde{\omega}_{i,t-1} \Delta z_{it} + \sum_{i \in C} \Delta z_{it} \Delta \tilde{\omega}_{it} + \\ &\sum_{i \in B} z_{it} \tilde{\omega}_{it} - \sum_{i \in D} z_{i,t-1} \tilde{\omega}_{i,t-1}.\end{aligned} \tag{8}$$

We stress that ΔP_t is fundamentally different from ΔI_t. It decomposes *all* of the output growth in an industry (not just the real productivity aggregate given by I_t).[14] Both of these decompositions

[14] ΔP_t may also be used to decompose the change in the level of *predicted* output growth, where the predictions are constructed using plant-level productivity estimates that are not the full plant-level output residual, but a residual instead that tries to control for factors the researcher wants to exclude from productivity growth (like measurement error, e.g.). See footnote 9.

satisfy the three conditions described above.

The first term in each case measures what we have referred to as the real productivity effect. It captures changes in plant-level productivity, weighting these changes with the initial period input share or input index. If no plant becomes more productive, this term would always be zero. If every plant becomes more productive, it is necessarily positive. In the case of (8), where the input index replaces the share, this term is the same as (5), and thus has the historical interpretation for productivity growth (i.e. change in aggregate output due to productivity growth holding inputs constant).

The second term captures growth due to the reallocation of inputs. Again, the specific interpretation can differ across the two decompositions. In (7), if more productive plants in the base period have a higher share of the (e.g. labor) inputs allocated to the industry, this term increases. Similarly, in (8), holding the aggregate input levels in an industry constant, if more productive plants in the base period use more inputs in the final period, this term increases. However, if industry output increases because there is input deepening in an industry (say, from other industries), this term will generally be positive for (8), while (7) is normalized so these do not show up in the decomposition.[15] These are the reallocation explanations for increases in aggregate productivity.

The third term is the product of the change in the input index and the change in productivity. As such, it contains both real productivity changes and rationalization effects. It is a measure of just how tied together the real productivity and rationalization cases are. To the extent that it is not small relative to the other terms, one cannot separately consider real productivity and rationalization changes. We return to this point momentarily.

The fourth term measures the impact of entrants on changes in aggregate growth. The fifth term measures the contribution of exiters. The impact of *net* entry is given by the sum of the fourth and fifth terms and is ambiguously signed. For our purposes, it will be sufficient to examine the impact of net entry on aggregate productivity changes, so we define net entry at time t to be $NE_t = \sum_{i \in B} s_{it}\tilde{\omega}_{it} - \sum_{i \in D} s_{i,t-1}\tilde{\omega}_{i,t-1}$, and replace these terms with the shorthand notation in future decompositions.[16]

[15] For example, if all inputs in an industry doubled but there were no changes in productivity, (7) reports zero change for reallocation and (8) reports the value of the change in output arising because all industry inputs have doubled.

[16] In some cases, though, one might want to capture the idea that entry contributes positively only if the new plants are in some sense more productive than the incumbents. Analogously, exit would contribute positively is the exiters are less productive than the survivors. The decomposition does not allow one to measure these effects. Foster et al. (2001) show how to examine deviations from means to address these concerns.

Often in applied work not all four terms of this decomposition are reported. Instead, the typical approach reports three terms, "resolving" the ambiguity of the covariance term by assigning some fraction of it to the real productivity term, (say Φ), and the other (1 - Φ) to the reallocation term. For I_t this is given as:

$$\Delta I_t = \left(\sum_{i \in C} s_{i,t-1} \Delta \tilde{\omega}_{it} + \Phi \sum_{i \in C} \Delta s_{it} \Delta \tilde{\omega}_{it} \right) + \left(\sum_{i \in C} \tilde{\omega}_{i,t-1} \Delta s_{it} + (1 - \Phi) \sum_{i \in C} \Delta s_{it} \Delta \tilde{\omega}_{it} \right) + NE_t. \tag{9}$$

For example, if one combines the covariance term with the rationalization term ($\Phi = 0$), the resulting decomposition simplifies to:

$$\Delta I_t = \sum_{i \in C} s_{i,t-1} \Delta \tilde{\omega}_{it} + \sum_{i \in C} \Delta s_{it} \tilde{\omega}_{it} + NE_t \tag{10}$$

This is probably the most commonly used decomposition of aggregate productivity.[17] The problem with this decomposition is that the second term, which is intended to capture rationalization, is a mix of the rationalization term and the covariance term, and hence is contaminated with real productivity changes (violating condition three).

Another commonly observed alternative is to fold the covariance term into the rationalization term (i.e. $\Phi = 1$):

$$\Delta I_t = \sum_{i \in C} s_{i,t} \Delta \tilde{\omega}_{it} + \sum_{i \in C} \Delta s_{it} \tilde{\omega}_{i,t-1} + NE_t \tag{11}$$

Now the real productivity term, $\sum_{i \in C} s_{i,t} \Delta \tilde{\omega}_{it}$, is a mix of the real productivity term and the covariance term, and hence is contaminated with rationalization effects (violating condition two).

Generally, it makes little sense to allocate the covariance term to either real productivity, reallocation, or both. Doing so reduces the available information (as adding variables together always does) and is thus a prescription for coming to false conclusions about the extent to which real productivity or rationalization drives industry growth.

Before turning to inference using this decomposition, we briefly describe one other approach sometimes used in the literature. Following Olley and Pakes (1996), one can decompose (4) in year t as:

$$O_t = \bar{\tilde{\omega}}_t + \sum_{i=1}^{N} \Delta s_{it} \Delta \tilde{\omega}_{it} \tag{12}$$

[17] It is used, for example, in Bailey et al. (1992), Bernard and Jensen (2001) and Tybout (1996).

where

$$\Delta s_{it} = s_{it} - \bar{s}_t, \quad \text{and} \quad \Delta\tilde{\omega}_{it} = \tilde{\omega}_{it} - \bar{\tilde{\omega}}_t$$

and $\bar{s}_t$ and $\bar{\omega}_t$ are the mean share and the *unweighted* productivity average respectively. The second term in (12) is the cross-sectional covariance between productivity and market share at time t.[18] The change in O_t over time then has two components, the change in the unweighted productivity average $\bar{\omega}_t$ and the change in the sample covariance between productivity and market share. As described shortly, the second term is useful for inference. However, the first term is an unsatisfactory measure of productivity growth because plant size plays *no* role in increases in the measured aggregate productivity growth.

5. Inference

What can be inferred about the long run growth prospects of an industry from the components of (7) or (8)? If the covariance term does not dominate the decomposition, real productivity growth and reallocation are separately identified from one another. Since sustainable growth prospects may be signaled by strong plant-level real productivity growth, these prospects show up in $\sum_{i\in C} s_{i,t-1}\Delta\tilde{\omega}_{it}$ (or $\sum_{i\in C} z_{i,t-1}\Delta\tilde{\omega}_{it}$). The relative role of real productivity vs. reallocation can be determined by comparing this term to $\sum_{i\in C}\Delta s_{it}\tilde{\omega}_{i,t-1}$ (or $\sum_{i\in C}\Delta z_{it}\tilde{\omega}_{i,t-1}$); the larger real productivity growth is, the better are prospects for sustained long-run growth. If the index from (8) is used, the units of the change can be interpreted as the impact in output or revenue terms of each growth component.

It is also possible (again, if the covariance term is small) to distinguish between the theory models described above using the decomposition. While real productivity growth is consistent with equilibriums in all of the theory models, reallocation is not. The stationary equilibriums of Jovanovic (1982), Hopenhayn and Rogerson (1993), and Melitz (2001) insist on no reallocation, thus requiring the reallocation term to be negligible. Finding that it is relatively large calls into question the appropriateness of using the equilibrium notions in any of these models for, say, counter-factual analysis.

If the covariance term is large, then as plants' productivity changes so too does their use of inputs. In this case, it is difficult to talk about the roles of real productivity and reallocation *separately* because they are so intertwined. However, we propose other ways of testing for the

[18] As it is written here, it is solely within period (i.e. cross-sectional), and thus violates condition one. Equation (12) informs the researcher whether, at a given point in time, it is larger plants that have a higher level of productivity. If used to examine productivity at only a given point in time, this decomposition simply cannot address questions concerning changes over time, and hence cannot address the real productivity and the rationalization stories.

existence of either the real productivity effects or the reallocation effects (not holding the other constant). These same methods can also help distinguish between the theoretical frameworks when the decomposition above cannot.

If there is any underlying real productivity growth in an industry, the distribution of productivity should be shifting over time. If the industry is in equilibrium and is well-described by a number of the theory models mentioned above, the productivity distribution should not change over time (up to a change in the mean that is common across firms). In both cases, Kolmogorov-Smirnov provide a distribution-free test that can answer whether two empirical distributions appear to have come from the same underlying population distribution.[19] The distribution-free nature of the test means that it imposes very little structure on the data when the testing is done.

This test is only available for comparing univariate distributions (no direct tests exist for comparing joint distributions). The equilibriums described above are often stationary in two state variables, productivity and capital or labor, suggesting that testing for differences in the marginal distribution of productivity may not be the most powerful approach; the joint distribution of state variables might change but leave the marginal distribution of productivity unchanged. Fortunately, many alternatives are available for testing. First, the distribution of productivities can be weighted by observed shares of capital or labor, and the new share-weighted empirical distribution functions tested for differences over time using Kolmogorov-Smirnov.[20] Second, changes in the second term from the index given by (12) reflect changes in the sample covariance between shares and productivity over time. A necessary condition that the joint distribution not change is that this sample covariance remain the same over time. Hence, computing this simple covariance for each period and examining the series informs the researcher about whether the joint distribution is changing.

A second set of analyses that sheds light on whether plant-level productivity is increasing over time tracks plant-level productivity observations directly. If plants are becoming more productive, plants should have higher productivity later in the sample than they had earlier in the sample. We implement three versions of this simple idea. First, we look at the sign of the year-to-year changes in plant productivity, asking what fraction are positive. Examining longer differences may be less subject to noise, so we also compute the fraction of plants that exhibit an increase in productivity

[19] The approach takes the maximal difference between two empirical distribution functions and applies a test statistic to it.

[20] Since both state variables enter into this test, it can uncover changes in the joint distribution when the marginals remain the same.

using long differences.[21] Our third implementation computes a plant's average productivity in the first half of its life (whatever that may be) and the plant's average productivity in the second half of its life. We then compute the fraction of plants for which the latter is larger than the former. Overall, if there is any evidence for real productivity growth (short or long term) it is likely to show up in at least one of these statistics. Because these tests are examining the share of changes that are positive, but not the magnitude of these changes, these tests are not directly useful for determining the magnitude of the real productivity effect. Rather, they are suggestive.

6. Data

The decompositions outlined in the previous section require very detailed plant-level data which has not been censored for entry and exit and which has a reasonable time-series dimension. The Chilean data set used in this study meets those requirements. The data span the period 1987 through 1996. Comparable data preceding this period (1979-86) have been used elsewhere and we refer the interested reader to those papers for a more detailed description of the data and, especially, a discussion of the economic changes in Chile spanning this sub-sample.[22] Here, we provide just an overview of the data.

The data are a manufacturing census covering all plants with at least ten employees. The data were originally provided by Chile's Instituto Nacional de Estadistica (INE). The structure of the data set is an unbalanced panel. There is information tracking plants over time and the data set includes plants that enter over the course of the sample period (births) as well as plants that exit (deaths.) Due to the way that the data are reported, we treat plants as firms, although there are certainly multi-plant plants in the sample.

In an attempt to keep the analysis manageable, we will focus on eight of the largest industries (excluding petroleum and refining.) We will work with industries at the 3-digit level. The industries (along with their ISIC codes) are Food Products (311), Textiles (321), Apparel (322), Wood Products (331), Printing and Publishing (342), Other Chemicals (352), Metals (381), and Non-electric Machinery (382). The data are observed annually and they include a measure of output, a measure of labor and capital inputs, and a measure of the intermediate inputs electricity and

[21] While the short differences examine plant-level productivity from one year to the next, the long differences examines the change in plant-level productivity where the difference is taken from the plant's first year in the sample to the plant's last year in the sample.

[22] A very detailed description of how the first eight annual samples were combined into a panel is found in Lui (1991). See also Lui (1993), Lui and Tybout (1996), Tybout, de Melo, and Corbo (1991), Levinsohn (1998), and most recently Levinsohn and Petrin (2003).

fuels. Real value-added is the real value of output adjusted for the real cost of all intermediate inputs.[23] Labor is the number of man-years hired for production, and plants distinguish between their blue- and white- collar workers. Electricity and fuels are measured in the real value of their volume consumed. Construction of the real value of capital is documented in Lui (1991).[24]

7. Estimation

Before one can investigate the micro-foundations of industry-level productivity, one must first estimate plant-level productivity. In this section we give a very brief overview of our preferred estimation strategy. We want to stress, though, that the investigation of the micro-foundations of productivity growth– the topic of this paper– does not depend on any particular estimation method. Whether one uses simple Ordinary Least Squares, revenue shares, or a state-of-the-art estimation technique, all result in a set of plant-level productivity measures that can then be aggregated into industry productivity. The details of how we estimate productivity in this paper are relegated to the Appendix, and readers not interested in them can safely skip to the "Results" section.

We measure productivity in the usual way, as the residual from an estimated production function. The exact details of our preferred approach is the topic of a related paper, "Estimating Production Functions using Inputs to Control for Unobservables," (Levinsohn and Petrin (2003)). There we show the importance of correcting for simultaneity even when many inputs are used in the estimation, including skilled and unskilled labor, capital, electricity, fuels, and materials. Here we focus on a relatively simple value-added production function, but one which is also likely to suffer from the same simultaneity problem.

Our production function has as arguments skilled and unskilled labor, capital, and the residual, and we write this function as

$$y_t = \beta_0 + \beta_k k_t + \beta_s l_t^s + \beta_u l_t^u + \omega_t + \eta_t, \tag{13}$$

where y_t is the log of output (measured as value-added) in year t, k_t is the log of the plant's capital stock, l_t^s is the log of skilled labor input, l_t^u is the log of the unskilled labor input, and the error is

[23] See the Appendix for more details on the construction of value-added.

[24] Essentially, it is a weighted average of the peso value of depreciated buildings, machinery, and vehicles, each of which is assumed to have a depreciation rate of 5%, 10%, and 20% respectively. No initial capital stock is reported for some plants, although investment is recorded. When possible, we used a capital series that was reported for a subsequent base year. For a small number of plants, capital stock is not reported in any year. We estimated a projected initial capital stock based on other reported plant observables for these plants. We then used the investment data to fill out the capital stock data.

assumed to be additive in two unobservables, ω_t and η_t. The component of plant productivity that is first-order Markov is given by ω_t. η_t is a mean zero error that may either be measurement error or a shock to productivity that is not transmitted and to which labor and other variable inputs do not respond. The key difference between ω_t and η_t is that the former is a state variable, and hence impacts the plant's decision rules, while the latter has no impact on the plant's decisions. A simultaneity problem exists if labor and capital respond to the plant-specific productivity term ω_t.[25]

The details of just how (13) is estimated are given in the Appendix. In an effort to make our estimation algorithm more readily available, exactly duplicable, and more user-friendly, the estimation algorithm has been adapted to run entirely in STATA and the program (in .do file format) will be available at either author's website.[26] We now turn to the results.

8. Results

Our goal in this section is to highlight issues that arise when examining the micro-foundations of industry productivity growth. We do this using recent data from manufacturing industries in Chile in order to bring home the point that the issues raised above are empirically relevant. We are neither trying to provide the definitive guide to productivity changes in Chilean manufacturing over the 1987-1996 period, nor argue that the potential problems discussed above *always* arise in every industry.[27] Rather, our aim is to illustrate that these problems *can* arise, and to provide some suggestions to protect against them and some tools that address them.

Before analyzing the micro-foundations of industry-level productivity, it is necessary to first estimate plant-level productivity. Table 1A (in the Appendix) gives the estimates of the production function parameters and their standard errors for the eight industries described in section 6. Electricity is used as our proxy and the standard errors are reported in parentheses. Details of the estimation algorithm and computation of the standard errors are provided in Levinsohn and Petrin

[25] Using (13), the endogeneity of inputs problem is readily illustrated. If the labor inputs chosen at time t respond to observed productivity ω_t, then the variable input choices in year t will be positively correlated with ω_t, leading to upwardly biased estimates of the elasticity of output with respect to labor. The capital coefficient may suffer from the same problem; capital is variable over time via depreciation and changes in investment. To make matters more difficult, capital and labor levels are highly correlated both within and across plants. Econometrically, this means that a positive bias in one coefficient can transmit a negative bias to the other coefficient since they are estimated simultaneously. See Levinsohn and Petrin (2003) for details.

[26] See either http://gsbwww.uchicago.edu/fac/amil.petrin/research/ or http://www.econ.lsa.umich.edu/~jamesl/LP. The algorithm is available as its own STATA command.

[27] See Bergoeing, Hernando, and Repetto (2003) for a more complete analysis of productivity growth in Chile between 1979-1999.

(2003).[28] The coefficients in Table 1A are almost always precisely estimated (the two exceptions concern β_k). We compute plant-level productivity using the parameter estimates from Table 1A as described in section 3.

The most popular measure for industry productivity (estimated using plant-level data) is defined as in (4), normalizing I_t to one year. Table 1 reports I_t/I_{1987}, where the average plant-level productivity is normalized to its value in 1987 for each of the eight industries, using labor shares as the plant-level weights. While year-to-year changes are negative in some instances, these are the exceptions, as industries appear to have experienced fairly steady productivity growth (with the exception of wood products (331) and perhaps apparel (322)).

For all the reasons discussed in Section 3, the results in Table 1 are – while standard – not easily interpreted economically. For example, if a policy maker had the choice between seeing productivity increase from 1.00 in 1987 to 1.36 in 1996 in textiles (321) or seeing it increase from 1.00 to 2.60 in machinery (382) over the same period, the results in Table 1 are simply non-informative. Indeed, from these numbers themselves there is no basis on which to know whether the impressive looking gain for ISIC 382 is large in any economically meaningful sense.

Table 2 provides ΔY_t, our alternative measure of industry productivity, which is based on (5). Again, we stress that it is different from (4). First, it is defined as productivity has historically been defined: the increase in output holding inputs constant. Second, the units have an economic interpretation (either in currency units or units of output). For each of the eight industries, ΔY_t is reported along with real value added (the percentage change in value added is given as $\Delta Y_t/Y_0$).[29] For ISIC 311, from 1995 to 1996 the change in industry value-added if plant-level inputs stayed constant but productivity changed by the estimated amount is 10.3 million Pesos. This figure, all by itself, conveys an economic magnitude that is readily compared to: industry-level value-added, productivity in other time periods for ISIC 311, productivity in other industries (i.e. across Table 2), and productivity in other countries (with some kind of cross-country deflator).

The numbers for ISIC 311 from 1995-96 also illustrate how different the productivity increase measured by ΔI_t can be from the historical definition of productivity given by ΔY_t. For example, using I_t, the measured productivity increase is about 6% (from 1.47 to 1.56). The true percentage increase is closer to 14% (10 million pesos divided by a base of 73 million pesos), almost 2.5 times

[28] The standard errors for the variable factors are from the first stage regression and the standard error on capital is the result of a bootstrap involving 50 draws.

[29] Because it is a first difference, there are no results for the first year of the sample, 1987.

larger. Thus, for many questions, the computation of ΔY_t, which is as easy to compute as (4), will be an important complement to I_t.

Equation (5) also facilitates cross-industry comparisons, something (4) simply cannot do. We provide two examples. Consider ISIC 311 from 1995-96 again. A quick overview of Table 2 shows that this increase is the *largest* annual increase for all eight industries between 1987-1996. While immediately apparent from Table 2, this is not at all apparent from looking at Table 1, where the change in the index from 1.47 to 1.56 in no way stands out from the rest of the numbers in the table.

Second, consider again the rather impressive looking performance of the non-electrical machinery industry (ISIC 382), where from 1987 to 1996 the I_t index increases by 160%. If one sums the annual changes in ΔY_t for ISIC 382, the cumulative productivity gain for the decade is about 10 Million Pesos. This long-term gain is less than the one year gain in ISIC 311 from 1995 to 1996, a fact completely lost on the reader when viewing Table 1, where the largest one year gain across all eight industries shows up as an inconspicuous 6% increase in productivity. By construction, ΔI_t does not keep track of the levels, and for some questions this can be quite important.

The Decompositions

We now turn to distinguishing between real productivity growth and growth derived from input reallocations. We use the decompositions of ΔI_t and ΔP_t discussed in section 4. Table 3A (ΔI_t) and Table 4A (ΔP_t) in the Appendix report the year-by-year decompositions for the largest industry in our sample, ISIC 311. The top row of Table 3 and Table 4 respectively reports the average values over time for each column (described momentarily) in Table 3A and 4A. Repeating this calculation for each industry provides the eight rows in Tables 3 and 4 respectively.

We start with Table 3 first, which decomposes ΔI_t, the standard index productivity measurement. The first column of Table 3 lists the industry ISIC code. Each of the next six columns reports the average across 1987-1996 for each industry for (in order): I_t, the average of ω's weighted by labor shares, ΔI_t, the annual change in I_t, and the components of the overall change for each industry from equation (7). These are, respectively, the reallocation term, the covariance term, the real productivity term, and the net entry term. Table 3 summarizes, for each industry, the numbers that would be in each industry's respective Table 3A.

The first row of Table 3 gives the results for ISIC 311. We discuss this row to illustrate the table's content. First, the average of I_t across ten years is 601.8, and the average ΔI_t is 28.0. Are these numbers large or small? It is not clear, for as noted above, these are averages (or changes in averages) of shift parameters that have no obvious economic magnitude.[30]

The overall change is comprised of four additive terms. Reallocation is about 1/4 the magnitude of the overall change (6.99). The covariance term is substantial and negative at -39.70, exceeding the overall change in absolute value. The real productivity term also exceeds the overall growth term, at 48.54. Finally, net entry on average contributes positively to changes in productivity. Again, the magnitudes have no direct interpretation, although their sizes relative to the overall change may help to give some suggestion of which term dominates the growth.

Table 4 reports the decomposition for ΔP_t. Here, by construction, all the figures are in the units of currency. The overall change in output is, on average, an increase of 4.5 million pesos a year, on a base average value added of almost 70 million pesos a year. This increase can arise because plants are producing more output holding inputs constant, because inputs are being shifted from less productive to more productive plants, or because total industry inputs are increasing over time. In ISIC 311 the reallocation component is equal to 2.1 million pesos, the covariance term is equal to -3.1 million pesos, the real productivity term is equal to 4.4 million pesos, and net entry contributes 1.1 million pesos.

An important point raised in section 4 regarding these decompositions - which applies to both Table 3 and Table 4 - concerns the inability to interpret decompositions that "fold" the covariance term into either the real productivity term, the reallocation term, or some mix of the two. Consider, for example, not reporting the covariance term separately for ISIC 311. If one adds the covariance term to the reallocation term, then on average reallocation is negative and real productivity and net entry account for all growth. If, on the other hand, one adds the covariance term to the real productivity term, then the reallocation term is twice the magnitude of the real productivity term. In fact, for half of the industries from Tables 3 and 4, where the covariance term is folded *reverses* the conclusion about what drives productivity growth in the industry for the 1987-1996 period. This illustrates the importance of separating the covariance term in every case.

The key messages about the flavor of productivity growth in Chile that come from Table 4 (or Table 3) are as follows. First, real productivity plays a major role in growth in some industries. In ISIC 311, 321, 342, 352, 381, and 382, even if all of the (negative) covariance term is added to the real productivity term, it is still positive and large. Second, reallocation also plays a role, although perhaps less decisively so, as even after attributing all of the negative covariance term

[30] The production technology from (1) and the plant level inputs must be taken into account.

to reallocation in ISIC 322, 342, 352, and 381, it still has a major role for growth. Third, the covariance term has a counter-intuitive sign in half the industries. A negative covariance means that, averaged across the decade, plants with falling productivities are growing in size. Finally, net entry contributes positively on average to industry productivity in 7 of the 8 industries.

Since the decompositions are not dispositive on the source of productivity growth in some of these cases, we turn to the auxiliary tests we suggested in section 5. We conclude the results section by illustrating these tests in particular instances. Our point is not that these tests are always going to be needed. Rather, in some cases, they are informative. We first provide two examples of how one might use the Kolmogorov-Smirnov test, which asks whether two empirical distributions could have been drawn from the same underlying population distribution. We then illustrate our other auxiliary tests.

Consider the decomposition of industry productivity for ISIC 381 as reported in Table 4. The practitioner may wonder whether there has been real productivity growth. The term capturing real productivity growth averages 737 while the covariance term averages -629. If one folded the covariance term into the real productivity term, the resulting sum is positive but not striking, making it difficult from this decomposition to assess whether real productivity is present.

One alternative is the Kolmogorov-Smirnov test, which can be used to compare the distribution of productivity in 1987 to that in 1996, and see if they appear to be similar. One can make the comparison for all plants, or for just the plants that exist in both 1987 and 1996 (i.e. holding net entry constant, as in the decomposition).[31] Looking just at plants that existed in both periods, Figure 1 compares the empirical cumulative distribution of plant-level productivities. The distribution for 1987 lies to the left of that for 1996, and the Kolmogorov-Smirnov test confirms these distributions are different beyond the one percent significance level. (The p-value to three decimal places for the K-S test is 0.000.) Thus, even when the decomposition is not dispositive, this test can provide evidence for or against productivity growth.

Another test is to compare two consecutive years and then analyze whether there has been real productivity growth. We do this for the transition from 1993 to 1994 for ISIC 311 (as reported in the Appendix in Table 4A.) The real productivity term in this instance was 4,514 (thousand pesos) while the covariance term was 4,318, again frustrating any inference on productivity growth. Using

[31] While this test can be more powerful than using the decomposition when looking for productivity growth, there are two dimensions along which it is weak. First, by comparing the first and last years, annual changes are compounded, so we are really asking whether the real productivity term, compounded over a decade, is always about zero. Second, the test is only looking at the cumulative distributions and does not in any way control for exchangability within the distribution.

Kolmogorov-Smirnov to test whether the distribution of plant-level productivities might have been drawn from the same underlying distribution gives a test statistic with a p-value of .01 indicating that the distribution of plant-level productivities in 1994 is different from 1993, that is, productivity changed.

For many models of economic growth, plant-level productivity growth is crucial. Table 5 reports results from three related tests that examine plant-level productivity growth directly. These tests are not as conducive to formal hypothesis testing as the Kolmogorov-Smirnov test, but neither do they suffer from the "exchangability" problem, where two identical distributions of productivities over time may be the result of lots of plant-level productivity changes but with every plant "exchanging" their productivity level with another plant so the two distributions look identical. The K-S test statistic is unchanged if two plants simply swap positions in the cumulative distribution in a given year, and thus cannot distinguish between these types of stories.

The tests reported in Table 5 examine productivity changes for each and every plant individually and so they speak directly to the issue of plant-level productivity increases. These tests, unlike the decomposition, are of a binary nature. They measure whether productivity increased at the plant level, but they do not account for the magnitude of that increase (or decrease.) Test 1 computes the fraction of plants that report an increase in year-to-year productivity. Test 2 compares the productivity of a plant in its first year and in its last year. The test statistic reports the fraction of plants for which productivity is higher in the last year. Test 3 takes the lifespan of the plant (whatever that may be) and divides it into a first half and a second half. The average productivity of the plant in the first half of its life is compared to its average productivity in the second half of its life. The test statistic reports the fraction of plants for which productivity was on average higher in the second half. Each test examines plant-level productivity changes, although Test 2 is more likely to be noisy in the sense that it relies heavily on just two observations for each plant and if either is in some sense unrepresentative for that plant, the results may be sensitive to these outliers.

The overall message of Table 5 is that real productivity increases were pervasive. For every industry, for Tests 1 and 3 more than half of the plants experience a productivity increase. For Test 2 at least 47% of plants show productivity growth for every industry. When one compares Table 5 with the role of real productivity in the decomposition in Table 4, the results are broadly mutually supportive. That is, the industries with the larger proportion of plants reporting positive productivity changes are also those industries for which the real productivity term in the decomposition is larger.

9. Conclusions

With the increased availability of plant-level data sets, the introduction of clever plant-level models of industry evolution and economic growth, and new econometric methods for estimating productivity, research on productivity has burgeoned. This paper sounds a cautionary note. On the critical side, we note that traditional measures of industry productivity are hard to interpret, that oft-used decompositions of industry productivity are potentially misleading, and that figuring out what underlies industry productivity growth is important if one is to link the econometric research with the existing theoretical literature. While a critical review is by itself informative, we provide some responses. We propose an alternative measure of industry productivity. We also propose some auxiliary tests to supplement the usual industry productivity decompositions.

Each of the points made in the paper are illustrated using recent Chilean data. That results suggests that real productivity growth has been significant in Chile over the past decade, that the supplemental tests we propose do what they are intended to do, and that our alternative measure of industry productivity tracks the traditional measure reasonably well (but with the advantage of being more readily interpretable).

References

Abowd, J., Kramarz, F., and Margolis, D. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–334.

Bailey, M. N., Hulten, C., and Campbell, D. (1992). Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity (Microeconomics)*, 187–249.

Bergoeing, R., Hernando, A., and Repetto, A. (2003). Productivity dynamics in chile. Harvard Mimeo.

Bernard, A., Eaton, J., Jenson, B., and Kortum, S. (2000). Plants and productivity in international trade. NBER Working Paper No. 7688.

Bernard, A., and Jensen, B. (1999). Exceptional exporter performance: Cause, effect, or both?. *Journal of International Economics*, *47*, 1–25.

Bernard, A., and Jensen, B. (2001). Exporting and productivity: The importance of reallocation. Dartmouth College.

Biesebroeck, J. V. (2001). Productivity dynamics with technology choice: An application to automobile assembly. Department of Economics, University of Toronto.

Eaton, J., Kortum, S., and Kramarz, F. (2001). An anatomy of international trade: Evidence from french firms. Boston University, Department of Economics.

Ericson, R., and Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, *62*(1), 53–82.

Foster, L., Haltiwanger, J., and Krizan, C. (2001). *New Developments in Productivity Analysis*, chap. Aggregate Productivity Growth: Lessons from Microeconomic Evidence, pp. 303–372. University of Chicago Press for the National Bureau of Economic Research.

Hopenhayn, H., and Rogerson, R. (1993). Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*, *10*(5), 915–938.

Hulten, C. (2001). *New Developments in Productivity Analysis, "Total Factor Productivity: A Short Biography"*, chap. 1, pp. 1–47. Universityof Chicago Press for the National Bureau of Economic Research.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, *50*(3), 649–670.

Levinsohn, J. (1998). Employment responses to international liberalization in Chile. Forthcoming in the Journal of International Economics.

Levinsohn, J., and Petrin, A. (1999). When industries become more productive, do firms? investigating productivity dynamics. NBER Working Paper 6893.

Levinsohn, J., and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, *70*(243), 317–342.

Lui, L. (1991). *Entry-Exit and Productivity Changes: An Empirical Analysis of Efficiency Frontiers.* Ph.D. thesis, University of Michigan.

Lui, L. (1993). Entry, exit, and learning in the chilean manufacturing sector. *Journal of Development Economics*, *42*, 217–242.

Lui, L., and Tybout, J. (1996). Productivity growth in Colombia and Chile: Panel-based evidence on the role of entry, exit and learning. In Roberts, M., and Tybout, J. (Eds.), *Producer Heterogeneity and Performance in the Semi-Industrialized Countries*, chap. 4. World Bank.

Melitz, M. (2001). The impact of trade on intra-industry reallocations and aggregate industry productivity. Department of Economics Harvard University.

Olley, S., and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*(6), 1263–1298.

Roberts, M., and Tybout, J. (1996). *Industrial Evolution in Developing Countries.* Oxford University Press for The World Bank, New York.

Tybout, J. (1996). Heterogeneity and productivity growth: Assessing the evidence. In Roberts, M., and Tybout, J. (Eds.), *Industrial Evolution in Developing Countries*, chap. 3. World Bank.

Tybout, J., de Melo, J., and Corbo, V. (1991). The effects of trade reforms on scale and technical efficiency: New evidence from Chile. *Journal of International Economics*, *31*, 231–250.

Appendix

Construction of Value Added

The measure of real value added that we use is constructed by subtracting the nominal value of raw materials, electricity, and fuels from nominal total sales revenues and deflating by the 3 digit industry level price deflator constructed by the Banco Central de Chile. We express value added as follows:

$$VA_{it} = (TSR_{it} - RM_{it} - E_{it} - \sum_j F_{ijt})/P_{OUTPUT,t}$$

where i indexes plants, VA is real value added, TSR is total sales revenues, RM is raw materials, E and F are value of electricity and fuels purchased, and j indexes the different fuels recorded in the census. The industry price deflator is given as $P_{OUTPUT,t} = p_{OUTPUT,t}/p_{80,t}$.

Zero Investment and the Electricity Proxy

The methods of Olley and Pakes (1996) revolve quite centrally around the investment decision and investment data. As discussed earlier, the estimation routine relies on investment being strictly increasing in productivity so the investment function can be inverted to proxy for productivity. They use data from the U.S. Census of Manufacturers, and they find that 8% of plant/year observations are reported to be zero. This feature of the data suggests that invertibility of $i(\cdot)$ fails at zero investment. Therefore, using investment as a proxy will require truncation of all observations at zero investment.

In the Chilean data, about one-third of the plant/year observations are reported to have zero investment. We are hesitant to truncate all of these observations from our estimation procedure. However, since we find that data on electricity is almost always reported at non-zero levels, we avoid this potential problem by relying on electricity as a proxy.

Estimation of the Production Function

Our estimator extends the idea in Olley and Pakes (1996). They use investment as a proxy for productivity, assuming that it is strictly increasing in productivity for positive levels of investment. We assume the intermediate input m is a strictly increasing function of ω. That is:

$$m_t = m_t(\omega_t, k_t), \tag{14}$$

and we then invert (14) and express the unobservable productivity as a function of the intermediate input and capital, or

$$\omega_t = h_t(m_t, k_t). \tag{15}$$

This inversion plays a very important role, since it permits us to control for ω_t. To see how this is done, substitute (15) into (13) to obtain:

$$y_t = \beta_s l_t^s + \beta_u l_t^u + \phi_t(m_t, k_t) + \eta_t, \tag{16}$$

where,

$$\phi_t(m_t, k_t) = \beta_0 + \beta_k k_t + h_t(m_t, k_t). \tag{17}$$

(16) is partially linear; it is linear in skilled and unskilled labor, and non-linear in the intermediate input and capital. We use data on electricity usage for the intermediate input m_t.[32] We

[32] See Levinsohn and Petrin (2003) and Levinsohn and Petrin (1999), where we show that results for production function estimates are robust across other intermediate inputs, including fuels and materials.

proceed by regressing y_t on l_t^u, l_t^s, and a third order polynomial in electricity (m_t) and k_t, i.e. we use a polynomial series to approximate the function $\phi_t(m_t, k_t)$.[33] We also allow for different functions $\phi_t(\cdot)$ depending on whether manufacturing growth is especially strong or weak.[34] Thus the first stage is as simple as OLS, and it yields estimates of β_l^u and β_s^u which are not contaminated by labor's responsiveness to the current period's productivity term; including $\phi_t(\cdot)$ controls for the correlation between labor and the error term.

We now describe how β_k is identified. From (17), capital appears twice in the equation and thus β_k is not identified without some further restriction. Next period's output is written as

$$y_{t+1} = \beta_0 + \beta_s l_{t+1}^s + \beta_u l_{t+1}^u + \beta_k k_{t+1} + \omega_{t+1} + \eta_{t+1}. \tag{18}$$

Define the function $g(\omega_t)$ as

$$g(\omega_t) = \beta_0 + E[\omega_{t+1}|\omega_t].$$

The function $g(\omega_t)$ gives, up to an additive constant, the expectation of next period's productivity, ω_{t+1}, conditional on this period's productivity shock. We can rewrite $\omega_{t+1} = E[\omega_{t+1}|\omega_t] + \xi_{t+1}$, where ξ_{t+1} is the innovation in productivity. The important identification assumption for capital is that k_{t+1} does not respond to this innovation (although it can freely covary $E[\omega_{t+1}|\omega_t]$). In practice, we estimate $g(\omega_t)$ non-parametrically, substituting it into (18) to provide the population moment

$$\begin{aligned} &E[y_{t+1} - \beta_s l_{t+1}^s - \beta_u l_{t+1}^u - \beta_k k_{t+1} - g(\omega_t)|k_{t+1}] = \\ &E[\xi_{t+1} + \eta_{t+1}|k_{t+1}] = 0. \end{aligned} \tag{19}$$

This moment identifies β_k.[35]

The second stage of the estimation uses $\hat{\beta}_l^u$, $\hat{\beta}_s^u$, and $\hat{\phi}_t(\cdot)$ to construct the sample analog to the moment restriction from (19) that identifies the capital coefficient. Given $\hat{\beta}_l^u$, $\hat{\beta}_s^u$, and $\hat{\phi}_t(\cdot)$, and any candidate value for β_k, say β_k^*, we can estimate the function $g(\omega_t)$ using a polynomial approximation with argument $\hat{\omega}_t(\beta_k^*) = \hat{\phi}_t(\cdot) - \beta_k^* k_t$. Alternatively, for any candidate value β_k^* we can compute the residual

$$[\xi_{i,t+1} + \eta_{i,t+1}](\beta_k^*)$$

for any plant i at time t (see equation (18).) We then use a non-linear least squares routine to locate the minimizer $\hat{\beta}_k$ which solves

$$min_\beta \sum_i \sum_{t=T_{i0}}^{T_{i1}} \left([\xi_{i,t+1} + \eta_{i,t+1}](\beta)\right)^2,$$

where T_{i0} and T_{i1} index the second and last period a plant is observed.

---

[33] In this and all future polynomial series approximations we experimented with a fourth order expansion and found that it had a negligible effect on our final parameter estimates.

[34] Hence, we allow for three different $\phi_t(\cdot)$ functions– one for normal times, one for expansionary periods and one for slow periods. Expansionary years include 1984, 1986, 1988, 1992, and 1995. Trough years include 1982, 1985, 1987, 1990, 1993, and 1996. The remaining years are the intermediate or "normal" years.

[35] It is perhaps helpful to note in less technical terms what this moment condition represents. The expectation of output less inputs equals the error, or productivity plus another additive independent error. This error cannot be used as the basis for a moment condition that will identify β_k, since productivity is not orthogonal to capital. We *can* solve for an error term, $(\xi_{t+1} + \eta_{t+1})$, that is uncorrelated with capital by conditioning out the expectation of ω_{t+1}. It is the inclusion of the function $g(\omega_t)$ which controls for this expectation and allows for identification of the capital coefficient (via the restriction from (19).)

TABLE 1

The Traditional Industry Productivity Measure (I_t/I_{1987})

(Reported Annually by ISIC Code)

Year	311	321	322	331
1987	1.00	1.00	1.00	1.00
1988	1.02	1.14	1.07	1.32
1989	1.03	1.12	1.06	1.27
1990	1.04	1.08	1.07	1.33
1991	1.17	1.16	1.20	1.37
1992	1.31	1.17	1.32	1.40
1993	1.39	1.17	1.31	0.96
1994	1.46	1.12	1.14	0.95
1995	1.47	1.20	1.15	0.91
1996	1.56	1.36	1.07	0.99

	342	352	381	382
1987	1.00	1.00	1.00	1.00
1988	1.30	1.03	1.19	1.33
1989	1.22	1.05	1.21	1.18
1990	1.31	1.27	1.30	1.17
1991	1.38	1.44	1.21	1.24
1992	1.53	1.50	1.32	1.46
1993	1.51	1.47	1.57	2.00
1994	1.57	1.58	1.46	2.32
1995	1.89	1.78	1.49	2.64
1996	1.71	1.88	1.65	2.60

TABLE 2

An Alternative Measure of Industry Productivity (ΔY_t)

(Thousands of Chilean Pesos))

	ISIC 311		ISIC 321		ISIC 322		ISIC 331	
Year	ΔY_t	Value Added	ΔY_t	Value Added	ΔY_t	Value Added	ΔY_t	Value Added
1988	586	39500	1354	9514	420	4782	1903	8790
1989	2186	42670	-232	10884	-56	5409	307	9020
1990	1926	44590	-60	10820	-18	5754	698	9033
1991	5207	50317	778	11635	748	6886	324	9643
1992	5208	58466	131	12119	407	7945	13	10956
1993	8086	63721	225	11582	625	8265	-2611	9182
1994	4514	68353	-875	10632	-1484	7567	-259	9464
1995	1674	73771	1017	11109	65	7330	-650	9177
1996	10301	77327	1208	11149	-523	6314	136	11389

	ISIC 342		ISIC 352		ISIC 381		ISIC 382	
	ΔY_t	Value Added	ΔY_t	Value Added	ΔY_t	Value Added	ΔY_t	Value Added
1988	997	5698	365	13644	1584	8830	608	3708
1989	-530	5988	288	15509	461	10793	-294	5216
1990	786	7087	3191	18748	422	12739	-149	5275
1991	752	8189	2684	22285	-1322	11420	154	5952
1992	859	9527	990	23293	1129	13944	1047	7036
1993	12	10722	1224	24784	1598	16861	2020	8903
1994	400	11894	437	26527	-1010	17113	1291	10979
1995	2203	14484	4313	30052	1430	18310	1082	12986
1996	-403	12336	99	34075	2018	19193	820	10144

TABLE 3
The Average (over time) of the Components
to Changes in Industry Productivity (ΔI_t)

Industry	Weighted Average of ω's	Overall Change	Realloc.	Covariance	"Real"	Net Entry
311	601.8089	28.0598	6.9902	-39.7072	48.5485	12.2282
321	888.3943	25.3024	12.3597	-4.5424	22.2532	-4.7681
322	130.9226	0.8368	1.4097	-3.1257	0.1100	2.4427
331	366.9978	-0.5152	-6.7500	-5.3640	-6.1588	17.7576
342	377.4156	17.9154	-0.1943	-1.7821	17.7115	2.1804
352	767.9177	47.2416	4.7744	-6.3637	45.7360	3.0949
381	120.8833	6.5994	1.3320	-3.3908	5.7349	2.9233
382	492.5022	53.3213	3.8211	-2.7700	45.7827	6.4874

TABLE 4
The Average (over time) of the Components
to Changes in Industry Productivity (ΔP_t)
(Thousands of Chilean Pesos))

Industry	Value Added	Overall Change in P_t	Realloc.	Covariance	"Real"	Net Entry
311	58072.64	4551.57	2102.12	-3174.11	4489.56	1134.00
321	11069.32	189.93	131.76	-267.66	384.92	-59.09
322	6774.52	241.42	448.41	-298.52	15.61	75.92
331	9671.34	612.82	284.96	-235.34	-124.32	687.52
342	9787.19	892.49	445.61	-179.41	563.76	62.54
352	23693.77	2232.04	963.10	-299.52	1522.94	45.51
381	14842.64	1355.29	992.15	-629.16	737.89	254.40
382	8204.47	781.75	49.29	-170.42	796.29	106.58

TABLE 5

The Fraction of Plants with Positive Productivity Changes

Industry	Test 1	Test 2	Test 3
311	0.597	0.507	0.757
321	0.621	0.487	0.647
322	0.600	0.532	0.632
331	0.593	0.496	0.360
342	0.622	0.477	0.742
352	0.605	0.487	0.747
381	0.576	0.498	0.706
382	0.621	0.510	0.760

Notes: Test 1 computes the fraction of plants with a positive year-to-year productivity change. Test 2 computes the fraction of plants that have higher productivity in their last year than in their first year. Test 3 computes the fraction of plants that have a higher average productivity in the second half of their life than in the first half.

TABLE 1A

Production Function Estimates

1987-1996

Industry Code	β_U	β_S	β_K	No. Plants No. Obs.
311	.441	.282	.18	1896
	(.010)	(.009)	(.069)	12110
321	.339	.508	.08	562
	(.021)	(.018)	(.046)	3155
322	.465	.420	.225	556
	(.020)	(.019)	(.048)	2713
331	.453	.384	.155	566
	(.023)	(.020)	(.038)	3078
342	.282	.398	.265	290
	(.022)	(.023)	(.057)	1625
352	.205	.460	.245	282
	(.027)	(.025)	(.072)	1594
381	.402	.400	.310	739
	(.018)	(.017)	(.103)	3643
382	.412	.419	.160	360
	(.024)	(.022)	(.072)	1698

Notes: Standard errors are in parentheses. "No. Plants gives the number of plants in the industry while "No. Obs." gives the number of plant-year observations in the industry.

TABLE 3A

Components to Changes in Industry Productivity ΔI_t (ISIC 311)

Year	Weighted Average of ω's	Overall Change	Realloc.	Covariance	"Real"	Net Entry
1988	324.9294	41.3895	-6.7710	-2.6842	47.9874	2.8574
1989	306.6248	-44.4440	-10.9826	1.6023	-35.6601	0.5964
1990	328.9908	37.7259	4.8819	0.3694	32.5845	-0.1099
1991	344.7972	6.9671	-18.9892	-6.8970	28.6922	4.1611
1992	383.7214	22.5963	2.3500	-8.1196	42.2626	-13.8967
1993	378.6353	-10.5493	-24.9463	-1.0156	-2.9353	18.3478
1994	393.4658	18.3101	9.5262	1.2604	8.8894	-1.3658
1995	473.9157	83.7300	16.8364	-4.0739	70.1540	0.8135
1996	428.6923	2.7626	22.0936	3.1499	-29.2823	6.8015
Mean	377.4156	17.9154	-0.1943	-1.7821	17.7115	2.1804

Notes: Results are not reported for 1987 because terms in the decomposition are not defined for the first year.

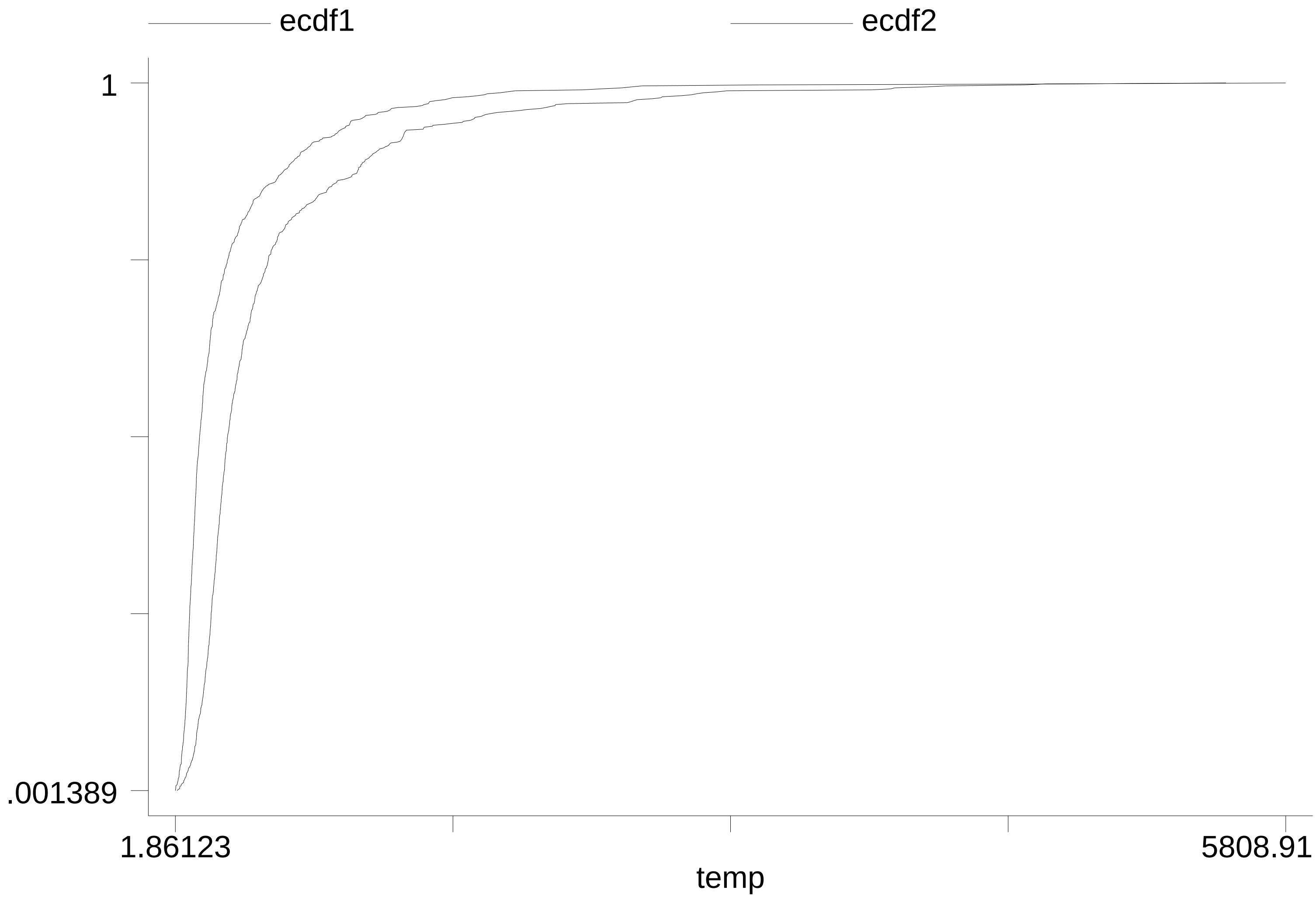
TABLE 4A

Components to Changes in Industry Productivity ΔP_t (ISIC 311)

(Thousands of Chilean Pesos))

Year	Value Added	Overall Change in P_t	Realloc.	Covariance	"Real"	Net Entry
1988	39500.48	3304.00	2857.97	-521.32	586.41	380.93
1989	42670.84	3188.36	1711.18	-1937.06	2186.12	1228.11
1990	44590.93	2065.69	524.69	-2095.49	1926.54	1709.95
1991	50317.30	5015.43	2706.69	-3845.84	5207.05	947.54
1992	58466.57	7164.36	2457.49	-2242.14	5208.18	1740.83
1993	63721.45	3897.64	3563.98	-6709.19	8086.15	-1043.30
1994	68353.45	4060.24	4969.69	-4318.91	4514.66	-1105.19
1995	73771.77	5385.00	2870.05	-1174.35	1674.31	2014.98
1996	77327.70	6472.88	-2871.19	-5313.11	10301.94	4355.25
Mean	58072.64	4551.57	2102.12	-3174.11	4489.56	1134.00

Notes: Results are not reported for 1987 because terms in the decomposition are not defined for the first year.



ecdf1
ecdf2
1
.001389
1.86123
5808.91
temp