

RESEARCH SEMINAR IN INTERNATIONAL ECONOMICS

Gerald R. Ford School of Public Policy
The University of Michigan
Ann Arbor, Michigan 48109-1220

Discussion Paper No. 552

**Measuring Aggregate Productivity Growth
Using Plant-Level Data**

Amil Petrin

University of Chicago

and

James Levinsohn

University of Michigan

December 2005

Recent RSIE Discussion Papers are available on the World Wide Web at:
<http://www.spp.umich.edu/rsie/workingpapers/wp.html>

NBER WORKING PAPER SERIES

MEASURING AGGREGATE PRODUCTIVITY
GROWTH USING PLANT-LEVEL DATA

Amil Petrin
James Levinsohn

Working Paper 11887
<http://www.nber.org/papers/w11887>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2005

The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2005 by Amil Petrin and James Levinsohn. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Contractual Versus Generic Outsourcing: The Role of Proximity
Amil Petrin and James Levinsohn
NBER Working Paper No. 11887
December 2005
JEL No. L0

ABSTRACT

We define productivity growth as the change in welfare that arises from additional output holding primary inputs constant. Using this traditional growth-accounting definition, we show that gains may arise because of plant-level technology shocks, and, in imperfectly competitive settings, from the reallocation of inputs across plants with differing markups and/or shadow values of primary inputs. With plant-level data, the alternative and most popular definition of productivity growth looks at the difference in the first moments of the productivity distribution. We show that this definition adds an additional term to the growth-accounting measure, which has been called “reallocation.” We show there is a very weak relationship between the two indexes in almost every 3-digit manufacturing industry in both Chile from 1987-1996 and Colombia from 1981-1991 - 49 in total - primarily because this “reallocation” term is large and volatile. We explore the theoretical reasons for this sharp divergence, in the process uncovering a number of previously unnoticed and unattractive features of the first-moment definition. For example, it is not tethered to any theoretical model, it is sensitive to measured units, and it can report positive productivity growth when welfare has fallen.

Amil Petrin
Graduate School of Business
University of Chicago
5807 Woodlawn Avenue
Chicago, IL 60637
amil.petrin@chicagogsb.edu

James Levinsohn
Department of Economics
University of Michigan
Ann Arbor, MI 48109-1220
and NBER
jamesl@umich.edu

Measuring Aggregate Productivity Growth Using Plant-level Data

Amil Petrin

University of Chicago, GSB

National Bureau of Economic Research

and

James Levinsohn

University of Michigan

National Bureau of Economic Research

First Version: December, 2003

Current Version: November, 2005

1. Introduction

Proper measurement of productivity growth is at the heart of a wide range of fields in economics. With the increasing availability of plant-level data, there is a large and rapidly growing body of research that estimates plant-level productivity and then aggregates to the industry level.¹ Typically, plant-level productivity is measured using the total factor productivity (TFP) residual ($\ln\omega$), computed as (log) output ($\ln Q$) minus the contribution of inputs ($\beta'\ln X$), or

$$\ln\omega = \ln Q - \beta'\ln X, \quad (1)$$

where (1) represents either the gross-output or the value-added production function. In this paper, we ask (and answer) “Can we aggregate these residuals in a way that yields a measure akin to welfare change?”

This paper grew out of initial work in “When Industries Become More Productive, Do Firms?: Investigating Productivity Dynamics” (NBER Working Paper 6893), and appeared in an earlier form as “On the Micro-Foundations of Productivity Growth.” We are grateful to the Russell Sage Foundation and the University of Chicago GSB for support and for helpful discussions from Bert Balk, Susanto Basu, Jeff Campbell, Steve Davis, Erwin Diewert, John Fernald, John Haltiwanger, Charles Hulten, Mitsukuni Nishida, Wendy Petropoulos, Mark Roberts, Johannes Van Biesebroeck, Alwyn Young, the NBER Productivity meetings, the Chicago Applied Micro Lunch, the Society for Economic Dynamics and the Econometric Society Meetings.

¹ The data sets used include U.S. data from the Longitudinal Research Database (LRD) of the U.S. Census, French data from the Déclarations Annuelles des Salaires (DAS) collected by INSEE (Institut National de la Statistique et des Etudes Economiques), and several plant-level manufacturing censuses from developing countries, to name but a few. Examples of papers using the U.S. data include Bailey, Hulten, and Campbell (1992), Olley and Pakes (1996), Bernard and Jensen (1999), Bernard, Eaton, Jensen, and Kortum (2003), and Foster, Haltiwanger, and Krizan (2001). Abowd, Kramarz, and Margolis (1999) and Eaton, Kortum, and Kramarz (2004) use the French data while the several papers in Roberts and Tybout (1996) use data from developing countries. A careful bibliography would include dozens of papers.

In many settings the answer is “yes” if we define aggregate productivity growth in historical, growth-accounting terms. Let i index the N plants in the economy, P_i denote i 's output price, Y_i equal the amount of i 's output going to final demand, and $d\omega_i^v$ equal the instantaneous change in final demand arising from plant-level technology shocks. Then, the resulting aggregate growth rate in value added from these shocks is

$$d\Omega = \frac{\sum_{i=1}^N P_i d\omega_i^v}{\sum_{i=1}^N P_i Y_i}. \quad (2)$$

In a competitive setting $d\Omega$ has a direct welfare interpretation, as it equals the change in society's ability to consume and invest holding primary inputs constant. $d\Omega$ continues to be tightly linked to welfare under many forms of imperfect competition (see Hulten (1978) and Basu and Fernald (2002), who use Domar (1961)). The welfare interpretation makes this measure useful for cost-benefit/policy analysis, and for the meta-analysis aggregating many studies, as it is readily comparable across time, industries, and countries.²

Multiplying $d\Omega$ by $\frac{Y_i}{Y_i}$ and rearranging yields the best-known formulation of (2):

$$d\Omega = \sum_{i=1}^N s_{v_i} d\ln\omega_i^v, \quad (3)$$

where $d\ln\omega_i^v$ denotes the instantaneous growth rate in the residual from the value-added production function and s_{v_i} denotes the plant's share of aggregate value added. Some approximation with discrete data from period $t - 1$ and t is needed, and two prominent alternatives exist. Tornqvist (1936), Theil (1967), Hulten (1973), Diewert (1976), Star and Hall (1976), Trivedi (1981), and Balk (2005), among others, have argued for multiplying the growth rate of productivity from $t - 1$ to t , estimated using $\Delta\ln\omega_{it}^v = \ln\omega_{it}^v - \ln\omega_{i,t-1}^v$, by the average share from the beginning and ending period:

$$\widehat{d\Omega} = \sum_i \frac{(s_{v_{it}} + s_{v_{i,t-1}})}{2} * \Delta\ln\omega_{it}^v, \quad (4)$$

and this yields the Tornqvist-Divisia quantity index.

The second approach - which is almost exclusively used - is some variant of the Bailey, Hulten, and Campbell (1992) index (the “BHC Index”):

$$BHC_t = \sum_i s_{v_{it}} \ln\omega_{it}^v - \sum_i s_{v_{i,t-1}} \ln\omega_{i,t-1}^v. \quad (5)$$

² Hulten (2001) provides a thoughtful and detailed history of this index.

One can characterize this index as a difference in first moments between period $t-1$ and t , with the “level” of aggregate productivity defined as the share-weighted average of log-level productivities. The widespread use of this index and many suggested variants can be found in the extensive literature review in Foster et al. (2001).

Our main empirical result is that these two measures differ significantly in their assessment of productivity growth on identical firms and in the aggregate. Using panel data for the 49 biggest manufacturing industries from Chile and Colombia, we show that industry-by-industry regressions of (4) on the BHC index yield slope estimates that are almost entirely less than 0.5, with most being in the range of 0.2 to 0.4, implying that BHC is typically larger in absolute value than (4) by several orders of magnitude. The R^2 s from these regressions generally range from between 0.2 to 0.4, indicating that these indexes are not linear transformations of one another.³

We explore the theoretical reasons for this sharp divergence. For firms existing in period $t-1$ and t BHC can be decomposed as

$$\begin{aligned} BHC_t &= \widehat{d\Omega} + \sum_i \frac{(\ln\omega_{it}^v + \ln\omega_{i,t-1}^v)}{2} * (s_{vit} - s_{vi,t-1}) \\ &= \widehat{d\Omega} + \sum_i \overline{\ln\omega_i^v} * \Delta s_{vi}, \end{aligned} \tag{6}$$

where $\overline{\ln\omega_i^v}$ is the average $\ln\omega_i^v$.⁴ Thus, BHC is equal to the growth-accounting measure *plus* $\sum_i \overline{\ln\omega_i^v} * \Delta s_{vi}$. Practitioners characterize this new term as reallocation’s role in productivity growth. A major drawback of this “reallocation” term is that it can be made arbitrarily large or small by appropriate choice of the (log) level of quantity. Empirically, as in our data, reallocation is often found to be large and volatile, dominating the BHC productivity index.

Closely related to this units problem - and perhaps fundamentally the root of all of the problems with BHC-type indexes - is the lack of a tight link to welfare. We are not aware of even one theoretical model of reallocation that yields $\sum_i \overline{\ln\omega_i^v} * \Delta s_{vi}$ as the welfare effect of reallocation. It

³ Overall, these findings are robust to a number of different methods of estimating plant-level TFP, including ordinary least squares, Solow’s approach, and a proxy approach.

⁴ See Griliches and Regev (1995), Fox (2003), and Diewert and Fox (2005), all of whom use Bennet (1920). The main variant of the BHC decomposition is given by

$$\sum_i \ln s_{vi,t-1} * (\ln\omega_{it}^v - \ln\omega_{i,t-1}^v) + \sum_i \ln\omega_{it}^v * (s_{vit} - s_{vi,t-1}).$$

which is similar to (6), and for this reason it will share the problems of (6) (plus one additional problem because it is not (6)).

is also easy to construct examples where the BHC index is *negatively* correlated with welfare, or where the BHC index reports productivity growth from reallocation when in fact there is none. For these reasons we argue it is important to have a theoretical model that motivates the use of any chosen productivity measure. This is especially true if one wants to undertake comparisons of the contribution of “real productivity” vs. “reallocation” to aggregate productivity growth, as is often done in the literature using the BHC index.

Real Productivity versus Reallocation Revisited

Given the difficulty of interpreting the economic meaning of the BHC reallocation term, we revisit the question of what it means to measure reallocation’s contribution to productivity growth. We focus on the case in which there is a fixed set of final goods, the most common setup for theoretical models of reallocation.⁵ Hall (1990) and Basu and Fernald (2002) show that reallocation effects only arise as we move away from a frictionless/perfectly-competitive world; they are the result of these imperfections. In a competitive setting, there is no increase in welfare that arises from simply reshuffling inputs across firms with different productivity levels. Here again, the BHC index fails, as its reallocation term will generally not be zero in the base case of perfect competition.

When markets are not perfectly competitive, there continues to be a direct link between welfare and the traditional growth-accounting definition of productivity if firms minimize costs and take inputs prices as given, and if a representative consumer model is a reasonable approximation to demand.⁶ With plant-level data, aggregate productivity growth so defined can be decomposed into five terms, one derived from the plant-level technology shocks and the other four from the reallocation of inputs across plants with different markups and/or shadow values of primary inputs.

We suggest also looking at the *change* in the productivity growth rate to further illuminate the underlying dynamics of reallocation and real productivity growth. If productivity growth is measured as in (4), we show that the change in growth decomposes into three terms. One term represents the change in aggregate growth from changes in plant-level productivity growth. A second term represents the change in growth from the reallocation of value added across plants with differing productivity levels. The final term represents a net entry effect. If a practitioner assumes perfect competition and wants to use growth-accounting techniques to measure the effect of input reallocation on welfare, we show that she *must* rely on decomposing changes in the productivity

⁵ See the review of the theoretical literature in Foster et al. (2001).

⁶ These conditions are satisfied by a large class of theoretical models that permit reallocation effects.

growth rate. In our data, the reallocation effect so defined is almost always present, economically important, reasonably stable, and almost always works to increase the growth rate in aggregate productivity, even in the instances where the growth rate in aggregate productivity falls.

Entry and Exit

We also revisit the issue of entry and exit, which is often given as the motivation for using the BHC index. We show that another attractive feature of the growth-accounting measure from (3) is that conceptually neither entry nor exit raise any difficulties. We argue that no compelling model has ever been put forward to motivate the BHC index treatment of entry and exit. We compare the treatment of entry and exit between the BHC measure and the growth-accounting measure and show that the two differ dramatically.

A missing values problem arises because we do not observe the exact time between $t - 1$ and t at which entry/exit occurs, nor do we observe the path of the entrant/exiter's productivity growth rate.⁷ We bound the error that arises from this missing information, and we provide estimators for the actual productivity contribution from entrants and exiters. In the process we show that the error in the aggregate from the missing data is small if these plants contribute only a small fraction to industry value added or have stagnant productivity growth— a proviso that is true in our annually observed data.

Sections 2-4 develop the economic model and the measurement methods. Section 5 addresses entry and exit and Section 6 introduces the change in productivity growth index. Readers interested only in the discussion of the widely used BHC index (and problems with it) as well as the associated empirical results can skip directly to Sections 7-9.

2. Aggregate Productivity Growth: Perfect Competition

We start with the competitive case with no frictions and a fixed set of final goods. This canonical case is a starting point for the more realistic models that allow for imperfections and/or frictions. In the competitive case, (3) exactly measures the shift in the social production possibilities frontier (PPF), as defined in Hulten (1978). It is equal to the change in welfare that arises as society's ability to consume and invest increases in response to the plant-level technology shocks. Let the constant returns-to-scale technology be given by $Q_i = F^i(M^i, X^i, t)$, where $(M^i = M_1^i, \dots, M_N^i)$ and $(X^i = X_1^i, \dots, X_K^i)$ denote respectively intermediate and primary inputs used at firm i . Let

⁷ For example, the Tornqvist-Divisia approximation requires observing data at both $t - 1$ and t .

P_i denote the output price and $P_i Y_i$ denote real final demand (or value added) from plant i , where $Y_i = Q_i - \sum_{j=1}^N M_j^i$ for good i . Aggregate final demand is given by $\sum_{i=1}^N P_i Y_i$.

From Solow (1957), total differentiation of $\ln Q_i$ and optimization together yield

$$d \ln Q_i = \sum_{j=1}^N \beta_j^i d \ln M_j^i + \sum_{k=1}^K \beta_k^i d \ln X_k^i + d \ln \omega_i = \sum_{j=1}^N s_{M_j^i} d \ln M_j^i + \sum_{k=1}^K s_{W_k^i} d \ln X_k^i + d \ln \omega_i \quad i = 1, \dots, N, \quad (7),$$

where $\beta_j^i = s_{M_j^i} = \frac{P_j M_j^i}{P_i Q_i}$ and $\beta_k^i = s_{W_k^i} = \frac{W_k X_k^i}{P_i Q_i}$, the elasticities of output with respect to the intermediate and primary inputs, which equal the revenue shares under perfect competition. Plant-level technical efficiency is

$$d \ln \omega_i = \frac{\partial F^i / \partial t}{F^i},$$

the well-known Solow residual. Lemma 1 shows that aggregate productivity growth $d\Omega$ is given by:

$$d\Omega = \sum_{i=1}^N \frac{P_i Q_i}{\sum_i P_i Y_i} d \ln \omega_i. \quad (8)$$

Hulten (1978) proves (8) for sector-level production functions (the type of data available in the 1970's) and Lemma 1 restates his argument when i indexes plants instead of sectors.

Lemma 1

Let $(Y = Y_1, \dots, Y_N)$, $(M = M_1, \dots, M_N)$, and $(P = 1, \dots, P_N)$ denote vectors of real final demand, intermediate input, and normalized output prices for the N plants in the economy. Let $(X = X_1, \dots, X_K)$ and $(W = W_1, \dots, W_K)$ denote vectors of primary inputs and their factor prices. Define the social production possibilities frontier implicitly using $F(Y, X, t) = 0$. Assume 1) $F(\cdot)$ is continuously differentiable and homogeneous of degree zero in (Y, X) , 2) that the economy is in a competitive equilibrium, and 3) the technology of each firm is characterized by constant returns to scale $Q_i = F^i(M^i, X^i, t)$ $i = 1, \dots, N$, with $(M^i = M_1^i, \dots, M_N^i)$ and $(X^i = X_1^i, \dots, X_K^i)$. Then the instantaneous shift in the social production possibilities frontier is the difference between the Divisia index of aggregate final demand and the Divisia index of total primary inputs:

$$d\Omega = \sum_{i=1}^N \frac{P_i Y_i}{\sum_{i=1}^N P_i Y_i} \frac{dY_i}{Y_i} - \sum_{k=1}^K \frac{W_k X_k}{\sum_{k=1}^K W_k X_k} \frac{dX_k}{X_k}, \quad (9)$$

and (9) is equal (8).

The proof follows Hulten directly and is provided in Appendix A. Alternatively,

$$d\Omega = \sum_{i=1}^N s_Y^i d \ln Y_i - \sum_{k=1}^K s_{X_k}^i d \ln X_k, \quad (10)$$

where $s_Y^i = \frac{P_i Y_i}{\sum_{i=1}^N P_i Y_i}$ and $s_{X_k}^i = \frac{W_k X_k}{\sum_{k=1}^K W_k X_k}$. The growth rate in aggregate final demand is given by the the sum across plants of their value-added share times their growth rate in final demand.

To hold primary input use constant, aggregate productivity growth deducts the aggregate growth rate of primary inputs, which is given by the sum across plants of their primary input share times their growth rate in primary inputs. Lemma 1 shows (8), (9), and (10) are all equal.

Equation (8) can be directly expressed in terms of the share of firm value added, denoted $s_{v_i} = \frac{P_i Y_i}{\sum_i P_i Y_i}$, and the residual from the value-added production technology. Let $s_{M^i} = \frac{\sum_{j=1}^N P_j M_j^i}{P_i Q_i}$ be the revenue share of intermediate inputs.

Lemma 2

$$d\Omega = \sum_{i=1}^N s_{v_i} d\ln\omega_i^v, \quad (11)$$

where $d\ln\omega_i^v = \frac{d\ln\omega_i}{(1-s_{M^i})}$.

See Appendix A for proof.

Domar (1961) first proposed these aggregation weights.⁸ In his words, they are defined in order to allow one to:

be free to take the economy apart, to aggregate one industry with another, to integrate final products with their inputs, and to reassemble the economy once more and possibly over different time units without affecting the magnitude of the Residual.

This greatly facilitates policy analysis because results are comparable across studies. Subsets of plants in the economy - such as industries or recent entrants/exiters - can be considered separately, and their relationship to the aggregate is directly known. For example, industry productivity can be aggregated with shares normalized to industry output, as it often is reported in studies, and the transformation to the aggregate economy is immediate via the scale factor given by the ratio of industry output to aggregate value added.

The competitive case illustrates what reallocation effects are *not*. When aggregate productivity growth is defined as the difference between aggregate value added and payments to primary inputs (as in (9)), primary inputs can be reallocated with no productivity/welfare reallocation effect. By the definition of productivity growth, the contribution to aggregate output of primary inputs is

⁸ (7)=(8)=(9)=(10) must account for the use of intermediate inputs in production. Mathematically, this means that the plant-level efficiency shocks must be grossed up by the ratio of plant gross output to aggregate value added. The story is as follows. When i 's output is part intermediate input, some of plant i 's output is used as input at other plants, so $d\ln\omega_i > 0$ leads both to an increase in final demand for i and to an increase in i 's intermediate deliveries. These new intermediate input deliveries are used in production at the plants to which they are delivered, where they increase output. This new output will both increase final demand and fulfill more intermediate deliveries elsewhere. The process continues. In the end, the greater the role of intermediate inputs used in the economy, the larger $\frac{\sum_{i=1}^N P_i Q_i}{\sum_i P_i Y_i}$ is, and the larger the impact of increases in plant-level technical efficiency on final demand.

deducted. In a competitive setting primary inputs are paid their marginal revenue product, which is equal to the contribution of the input to aggregate output. This is the *only* mechanism through which input reallocation enters (9), so it is the only way reallocation could possibly affect welfare (and it does not). Thus, any measure that reports a change in productivity when there is no welfare change - such as the widely used BHC measure - fundamentally differs from the traditional growth-accounting definition.

3. Aggregate Productivity Growth: Imperfect Competition

With imperfect competition, Hall (1990) shows that the productivity residual measures more than just the effects of technical progress on final aggregate demand. Basu and Fernald (2002) link the change in measured aggregate productivity with imperfect competition to the change in welfare, showing that input reallocations do lead to welfare change when we move to a world where there are markups and/or differences between firms in terms of their shadow values of primary inputs. We summarize Basu and Fernald's results in Lemmas 3 and 4.

For Lemma 3 we must define a general equilibrium framework. Suppose there are $N - 1$ consumption goods and an N th investment good. Consumption at time t of good j is denoted $C_{t,j}$ and aggregate value added is given by $Y_t = \sum_i P_{it}Y_{it}$. \bar{L} is the per-period endowment of labor, A_t are assets, P_{It} is the price of the investment good, P_{Lt} and P_{Kt} are the wage and rental price of capital, δ is the depreciation rate of capital, K , and Π_t are pure profits, which are rebated to consumers.⁹

Lemma 3

Suppose a large number of identical households choose consumption, investment, and labor to maximize lifetime utility:

$$\text{Max } U = \sum_{s=0}^{\infty} \beta^s u(C_{1,t+s}, C_{2,t+s}, \dots, C_{N-1,t+s}, \bar{L} - L_{t+s})$$

such that

$$A_{t+1} = A_t + P_{Lt}L_t + P_{Kt}K_t + (1 - \delta) * (P_{I,t+1} - P_{It})K_t + \Pi_t - \sum_{i=1}^{N-1} P_{i,t}C_{i,t}.$$

Let a temporary one-period shock change assets, consumption, and/or labor supply. Then a first-order approximation to the change in utility is proportional to

$$d\Omega_t = d\ln Y_t - \frac{P_{Kt}K_t}{Y_t} d\ln K_t - \frac{P_{Lt}L_t}{Y_t} d\ln L_t. \quad (12)$$

⁹ For simplicity we ignore private bonds, which in equilibrium will equal zero because arbitrage provides for an interest rate leading households to be indifferent between holding a bond and holding capital.

See Basu and Fernald (2002) for a proof in a more general setting. Lemma 3 shows that the change in welfare is proportional to the change in productivity, defined as the aggregate growth rate of value added minus the weighted aggregate growth rates of primary inputs, where the weights are their revenue share of value added. The result is a direct consequence of the envelope theorem; consumers are optimized prior to the shock, so their marginal rates of substitution between goods are equal to relative market prices, including the marginal rate of substitution between labor and leisure.¹⁰ The change in output net of inputs that arises from the shock - valued at market prices - provides a first order approximation to welfare (see Appendix A for a simple example).

Lemma 4 links (12) to plant-level data. In this setting, with potentially quasi-fixed inputs, Berndt and Fuss (1986) show that productivity must be measured using the shadow rental price of the input as opposed to the unit price.¹¹ Let $K = \sum_i K_i$ and $L = \sum_i L_i$. We denote P_{L^i} and P_{K^i} the shadow values of labor and capital respectively, and we define $P_L = \frac{\sum_i P_{L^i} L_i}{\sum_i L_i}$ and P_K analogously, as the “average” shadow values of primary inputs across firms. Value added is given by its standard definition:

$$d\ln Y_i = \frac{d\ln Q_i - s_{M^i} d\ln M_i}{1 - s_{M^i}} \quad (13),$$

which deducts the contribution of intermediates and then divides by $1 - s_{M^i}$ so that $d\ln Y_i = d\ln Q_i$ if $d\ln M_i = d\ln Q_i$. Equation (13) implies a value-added input index equal to

$$dX_i^v = \frac{s_{K^i}}{1 - s_{M^i}} d\ln K_i + \frac{s_{L^i}}{1 - s_{M^i}} d\ln L_i,$$

where s_{L^i} , s_{K^i} , and s_{M^i} denote the revenue shares. We define markups as

$$\mu_i = P_i / MC_i,$$

where MC_i denotes marginal cost. Finally, let the “value-added markups” equal

$$\mu_i^v = \mu_i \frac{1 - s_{M^i}}{1 - \mu_i s_{M^i}},$$

where μ_i^v is an increasing function of μ_i .

¹⁰ The marginal rates of substitution are not equal to the marginal rates of transformation because of markups.

¹¹ In a frictionless setting these two prices are equal.

Lemma 4

If firms minimize costs and take input prices as given, then plant-level value added can be written as

$$d\ln Y_i = \mu_i^v dX_i^v + (\mu_i^v - 1) \frac{s_{M^i}}{1 - s_{M^i}} (d\ln M^i - d\ln Q_i) + \frac{d\ln \omega_i}{1 - \mu_i s_{M^i}}. \quad (14)$$

The plant-level productivity residual is given as

$$\begin{aligned} d\ln \omega_i^v &= d\ln Y_i - dX_i^v \\ &= (\mu_i^v - 1) dX_i^v + (\mu_i^v - 1) \frac{s_{M^i}}{1 - s_{M^i}} (d\ln M^i - d\ln Q_i) + \frac{d\ln \omega_i}{1 - \mu_i s_{M^i}}. \end{aligned} \quad (15)$$

Aggregate productivity growth is equal to

$$\frac{1}{Y_t} * \sum_{i=1}^N [P_i Y_i * d\ln \omega_i^v + L_i * (P_{L^i} - P_L) d\ln L_i + K_i * (P_{K^i} - P_K) d\ln K_i]. \quad (16)$$

See Appendix A for proof. Perhaps most remarkable is that this set of assumptions does not require profit-maximization and is thus robust to any type of static or dynamic price-setting behavior, such as pricing policies derived from supergames.

If one substitutes (15) into (16), aggregate productivity growth is additively separable in five terms, three of which are contained in the plant-level productivity residual. From (15), $\frac{d\ln \omega_i}{1 - \mu_i s_{M^i}}$ is the only term related to the impact of technical progress on final demand. It resembles the perfectly competitive case except for the presence of the markup term in the denominator; the higher the markup, the larger the wedge between the value of output and its cost of production, and the more valuable to society the additional output generated by the shock to technical efficiency.

The other four terms all arise in response to the reallocation of inputs across plants. We review each in turn. The first term in (15) arises because only dX_i^v is deducted from $d\ln Y_i$ (instead of $\mu_i^v dX_i^v$). If inputs are reshuffled to plants with higher markups this term serves to increase the measured productivity residual. The second term arises because the productive contribution of intermediate inputs exceeds the revenue share because of the markup. Similar to the previous term, the revenue share is used to deduct intermediates from gross output in the move to value added. Thus, productivity increases as plants that use intermediate inputs more intensively increase their output. The final two terms, which are the last two terms in (16), relate to shadow values. For the two primary inputs labor and capital, aggregate productivity growth increases if there is a reallocation of inputs from plants with lower shadow values to plants with higher shadow values. Overall, any reallocation of inputs towards firms with higher markups and/or shadow values of capital or labor cause both productivity and welfare to increase.

Lemmas 3 and 4 together illustrate that technical efficiency is only one component of aggregate productivity growth when markets are not perfectly competitive. In the case when markets are competitive, μ_i and μ_i^v both equal one and all firms share a common shadow value for each primary inputs, so all reallocation terms in this case are zero, and welfare change is equal to $d\ln\omega_i^v = \frac{d\ln\omega_i}{1-s_{Mi}}$, the rate of productivity growth under perfect competition. Thus the residual from the imperfectly competitive case is in the limit - when prices converge to marginal cost and shadow values converge to unit prices - the productivity residual under perfect competition.

4. Estimating Productivity Growth

In this section we focus on estimation of aggregate productivity growth from period $t - 1$ to period t , as defined by

$$\Omega_{[t-1,t]} = \int_{t-1}^t \sum_i s_{v_{it}} d\ln\omega_{it}^v. \quad (17)$$

This is simply the integral of (3) over the period $t - 1$ to t . There are two steps to estimating (17). Plant-level productivity residuals must be recovered. Then they must be aggregated. We discuss each in turn.

Two prominent and complementary methods are frequently used to estimate plant-level productivity residuals. One method begins with an assumed functional form for the production function and then directly estimates its parameters.¹² The second approach uses the Solow (1957) insight that optimizing behavior implies observed revenue shares are a consistent estimator in the competitive setting for the elasticity of output with respect to any input. As shown in the proof of Lemma 4, in an imperfectly competitive setting, these shares consistently estimate the markup times the revenue share.

Data are observed discretely, requiring some discrete time approximation to the integral in (17). For productivity measurement the most popular discrete-time approximations are given by log-change indexes.¹³ In this case, $\ln\omega_{it}^v - \ln\omega_{i,t-1}^v$ is used as an approximation to the growth rate in value-added productivity over the time period. In the calculation of (17), a weight α_i is applied to the estimated average growth rate at each firm, and then aggregated across plants, yielding the approximation:

$$\sum_i \alpha_i (\ln\omega_{it}^v - \ln\omega_{i,t-1}^v).$$

¹² One can use ordinary least squares or one of many alternatives that attempt to address the simultaneity of input choices and productivity raised by Marschak and Andrews (1944). Alternatives include (for example) instrumental variables, fixed effects, and the proxy methods of Olley and Pakes (1996) and Levinsohn and Petrin (2003).

¹³ Sato (1987) in New Palgrave describes the properties of these approximations.

For α_i , Tornqvist (1936) advocates averaging beginning and ending period shares. In the context of the value-added residual $dln\omega_i^v$, the Tornqvist weight is given by

$$\alpha_i = \frac{s_{vit} + s_{v_{i,t-1}}}{2}.$$

In the context of $dln\omega_i$, the residual from the full production function,

$$\alpha_i = 1/2 * \left[\frac{P_{it}Q_{it}}{\sum_i P_{it}Y_{it}} + \frac{P_{i,t-1}Q_{i,t-1}}{\sum_i P_{i,t-1}Y_{i,t-1}} \right].$$

Theil (1967), Hulten (1973), Diewert (1976), Star and Hall (1976), Trivedi (1981), and Balk (2005), among others, have all argued for the Tornqvist approximation.¹⁴

Lemma 5

The Tornqvist approximation to the Divisia index is superlative (or without error) if the underlying function has the homogeneous translog form. With constant returns to scale the Tornqvist approximation to the Divisia index is uniquely superlative.

See Diewert (1976), who uses the quadratic approximation lemma to single out the Tornqvist approximation as most preferred. When the underlying function is not translog, the Tornqvist approximation remains attractive because the translog form provides a second-order approximation.

Trivedi develops approximation results for a more general class of functions.

Lemma 6

The error in the Tornqvist approximation to the Divisia index is on the order of the square of the length of the time interval.

See Trivedi (1981). Note that Paasche and/or Laspeyres log change approximations use $\alpha_i = s_{vit}$, the ending period share (Paasche), or $\alpha_i = s_{v_{i,t-1}}$, the beginning period share (Laspeyres). They will generally have more approximation error because they both ignore one of the two available pieces of information on the path from $t - 1$ to t of the share-weight. Table 4 reports estimated growth using log-changes with average share (Tornqvist) and base-period share, and shows they are quite different in our annual plant-level data (this is true because shares change in a substantial way). Lemmas 5 and 6 and our empirical findings are the reasons we prefer the Tornqvist approximation to the Divisia index.¹⁵

¹⁴ Readers (albeit perhaps a select few) may recognize the use of the average of the starting and ending share as Newton's trapezoid rule for approximating the area under a continuous curve.

¹⁵ Before moving on we note that we have not addressed a more fundamental (although perhaps less well-known) question. (17) is a line integral, so we must either observe the path followed between $t - 1$ and t , which we do not, or we must make do with observations at $t - 1$ and t , in which case the value of the line integral must be independent of the path taken. Hulten (1973) explores the necessary and sufficient conditions for (17) to satisfy path independence. Petrin (2005) discusses the implications of these conditions for using plant-level data. See also Balk (2003).

5. Accounting for Entry and Exit

One important issue that does arise with the move from industry- to plant-level data is that industries do not enter and exit, while plants do. Neither entry nor exit raises any conceptual difficulties for the aggregate productivity growth measure proposed in Lemma 1, when the set of final goods is fixed, as in Hulten (1978), Basu and Fernald (2002), and most dynamic general equilibrium models. However, new goods raise the standard questions with welfare measurement in the face of insufficient price variation, as discussed in Petrin (2002) and Goolsbee and Petrin (2004).¹⁶

Let $1_{i,Continue}$, $1_{i,Enter}$, and $1_{i,Exit}$ be indicator variables for whether firm i is a continuing firm, an entrant, or an exiter respectively for the period $t - 1$ to t . Firm i 's contribution to aggregate productivity growth from period $t - 1$ to t is then given as

$$1_{i,Continue} \int_{t-1}^t s_{vit} d\ln\omega_{it}^v + 1_{i,Exit} \int_{t-1}^{t^*} s_{vit} d\ln\omega_{it}^v + 1_{i,Enter} \int_{t^*}^t s_{vit} d\ln\omega_{it}^v, \quad (18)$$

where t^* denotes the time of entry or exit and $s_{vit^*} = 0$. Specifically, a continuing firm contributes $\int_{t-1}^t s_{vit} d\ln\omega_{it}^v$, an exiting firm contributes $\int_{t-1}^{t^*} s_{vit} d\ln\omega_{it}^v$, and an entrant contributes $\int_{t^*}^t s_{vit} d\ln\omega_{it}^v$.

Since observations on plants are only made at discrete intervals, for plants that enter and exit the choice of discrete approximation must be revisited because we do not have enough information to calculate the Tornqvist-Divisia quantity. In particular, while we know that $s_{vit^*} = 0$ at the time of entry/exit and from that time backward/forward, we do not observe the exact time between $t - 1$ and t at which entry/exit occurs, nor do we observe the path of i 's productivity growth rate during its existence in this period. This will be a problem for any measure that is based on following the path of plant-level productivity.

An advantage of the growth-accounting framework is that it tells us exactly how much error is added, and it suggests both bounds on the error and approximations to it. Before turning to these measurements, we note that many economic questions relate to the performance of entrants or exiters. Even in the worst case scenario in which all entrants (exiters) are truncated from the productivity calculation in their entering (exiting) year, one can compute these plants' contribution to aggregate productivity growth in the other years of their existence. Thus, for annual data, categories for recent entrants and/or upcoming exiters can be defined separate from continuing

¹⁶ Petrin (2005) addresses this question when plant-level data are observed and the change in welfare is the object of interest.

plants, their fraction of productivity growth being then directly comparable to continuing plants. *In fact, a principle feature of the growth-accounting measure is that we are free to group plants in any subaggregates that we desire without affecting the measure of aggregate productivity growth.*

Using (18), it is possible to bound the error that arises from this missing information. Let the contribution of all continuing plants to aggregate productivity growth be given by:

$$C = \sum_{i=1}^N 1_{i,Continue} \int_{t-1}^t s_{v_{it}} d\ln\omega_{it}^v.$$

Similarly, let exiters' contribution be

$$EX = \sum_{i=1}^N 1_{i,Exit} \int_{t-1}^{t^*} s_{v_{it}} d\ln\omega_{it}^v \quad (19)$$

and entrants

$$EN = \sum_{i=1}^N 1_{i,Enter} \int_{t^*}^t s_{v_{it}} d\ln\omega_{it}^v. \quad (20)$$

Then the true value for aggregate productivity growth over the period $[t-1,t]$ is bounded:

$$\Omega_{[t-1,t]} \in [C - |EX| - |EN|, C + |EX| + |EN|]. \quad (21)$$

The error in aggregate productivity growth from truncating entrants and exiters will be small if $|EX|$ and $|EN|$ are small. $|EX|$ and $|EN|$ are likely to be small when these plants contribute only a small fraction to industry value added or have stagnant productivity growth.

The fraction of value added that is attributable to these plants can be directly calculated from the data. In the two annual censuses from Chile and Colombia that we use in the empirical work, continuing plants account for between 94% and 98% of industry value added. When data are observed at 5 year intervals, the fraction of value added accounted for by entrants and exiters ranges from 10% to 40%. Thus, with annual data no additional correction may be necessary, but with data collected at 5 year intervals some approximation to the contribution to growth from entrants and exits may be desirable. We turn to this question next.

Approximations

We develop the approximations using the Tornqvist-Divisia measure as our estimation objective. There are two related issues for any approximation to an individual plant's contribution to $|EX|$ or $|EN|$: what is the appropriate estimator for i) the share-weight; and ii) the growth rate of productivity. A complete model of industry evolution may provide additional structure useful in

inferring the unobserved paths of technical efficiency and value added for entrants and exiters. We suggest estimators for the share-weight and productivity path when a full model is not specified.

We first focus on the share-weight and consider the case of exit; the case of entry is symmetric. Our goal is an estimate of the average share between time $t - 1$ and time t . We know for exiters that the share weight decreases (possibly monotonically) from the observed share at time $t - 1$ to zero some time after $t - 1$, and remains at zero until t . This means we know that the true Tornqvist weight lies somewhere on the interval $[0, \frac{s_{v_{i,t-1}}}{2}]$. Hence, $\frac{s_{v_{i,t-1}}}{2}$ provides an upper bound that approaches the Tornqvist weight as the time of exit approaches t , and we suggest using $\frac{s_{v_{i,t-1}}}{2}$ as the weight. The bound is also helpful because it can be used in a resampling procedure we describe momentarily that estimates the variability in the estimated values of $|EX|$ or $|EN|$.

Next we consider possible estimators of the average growth rate of productivity. For the exit case we want an estimate of $\ln\omega_{it^*} - \ln\omega_{i,t-1}$. The missing piece of information is $\ln\omega_{it^*}$, the productivity level at the time of exit. We suggest forecasting $\ln\omega_{it}$ using past values of $\ln\omega$ and any other relevant state variables. Given an estimate $\widehat{\ln\omega_{it}}$, one can use $\widehat{\ln\omega_{it}} - \ln\omega_{i,t-1}$ as the approximation to $\ln\omega_{it^*} - \ln\omega_{i,t-1}$. The approximation converges to the Tornqvist index as t^* approaches t and $\widehat{\ln\omega_{it}}$ converges to $\ln\omega_{it}$.

Some examples may be helpful. If productivity is posited as a first-order Markov process, a regression of $\ln\omega_{it}$ on previous log-levels of productivity provides an estimate of

$$E[\ln\omega_{it} | \ln\omega_{i,t-1}].$$

A second-order Markov process would include $\ln\omega_{i,t-2}$ in the conditioning set. If the other state variables are co-determined with productivity, then the regression might include these state variables. For example, a first order Markov process with state variables productivity and capital k would point to

$$E[\ln\omega_{it} | \ln\omega_{i,t-1}, k_{i,t-1}]$$

as the key equation of interest in forecasting the missing productivity level. In the case of entry, the conditional mean will use future outcomes to project the current state at the time of entry.

An important question in the construction of the estimator is what set of plants to use when estimating this process. If entrants and exiters are assumed to be similar to continuing plants, all plants can be pooled in an estimation routine. Alternatively, if the process for recent entrants and recent exiters is thought to be different, we could select on the appropriate set of plants and estimate on this subset.

Robustness analysis will be important, especially as the time interval between plant-level observations grows. Many resampling schemes are available that will illustrate the variability in $|\widehat{EX}|$ and $|\widehat{EN}|$. In the case of entry, random draws can be taken from the (conditional) distribution of $\ln\omega_{it}$ and from the observed interval $[0, \frac{s_{v_{i,t-1}}}{2}]$ to generate a range of possible estimates of the contribution to aggregate productivity growth of the entering firm i in period t . Repeating this with all entrants and aggregating provides an estimate of the variability introduced into aggregate productivity from not observing the path of the entrant. The estimator is easily modified for the exiters. Other sources of sampling error, such as that in the estimated equation for productivity, are easily incorporated into this resampling scheme in the usual way.

6. Decomposing Changes in Aggregate Productivity Growth Rates

In this section we show how the *change* in the productivity growth rate provides further information on the underlying dynamics of real productivity growth and reallocation. We suggest an approximation to the change in the growth rate that uses data from three juxtaposed time periods, which we denote as t , $t + 1$, and $t + 2$. Our estimated change in the growth rate is given by the difference in two “average-share log-change” indexes (i.e. Tornqvist-Divisia indexes), which using the value-added formulation gives:

$$\sum_{i=1}^{N_2} \frac{s_{v_{i,t+2}} + s_{v_{i,t+1}}}{2} (\ln\omega_{i,t+2}^v - \ln\omega_{i,t+1}^v) - \sum_{i=1}^{N_1} \frac{s_{v_{i,t+1}} + s_{v_{it}}}{2} (\ln\omega_{i,t+1}^v - \ln\omega_{it}^v). \quad (22)$$

Equation (22) decomposes into a “productivity” term, a “reallocation” term, and a “net entry” term. The set of plants that exist in t , $t + 1$, and $t + 2$ is denoted C (for continuers), and each of these plants contributes a productivity component and a reallocation component to the growth in aggregate productivity. Plants that exist only in t and $t + 1$, or only in $t + 1$ and $t + 2$, contribute to the overall change in (22), but only through the “net entry” term described below.

The productivity term is written as

$$\sum_{i \in C} \frac{s_{v_{i,t+2}} + 2 * s_{v_{i,t+1}} + s_{v_{it}}}{4} * \left(\ln \frac{\omega_{i,t+2}^v}{\omega_{i,t+1}^v} - \ln \frac{\omega_{i,t+1}^v}{\omega_{it}^v} \right).$$

Each continuer contributes the change in the rate of their productivity growth (the term on the right). The weight in the aggregation to the industry level is an average over the three time periods of the plant’s share in value added, where period $t + 1$ gets twice the weight of period t and period $t + 2$.

The second term is the reallocation term, given by:

$$\sum_{i \in C} \left(\ln \frac{\omega_{i,t+2}^v}{\omega_{i,t+1}^v} + \ln \frac{\omega_{i,t+1}^v}{\omega_{it}^v} \right) / 2 * (s_{v_{i,t+2}} - s_{vit}) / 2.$$

The term on the right gives the change in the share of value added from period t to period $t+2$. This multiplies the average rate of productivity growth over the periods for the plant (i.e. $\ln(\frac{\omega_{i,t+2}^v}{\omega_{it}^v})/2$).

The third term is the net entry term and is given as

$$\sum_{i \in t+1, t+2, \text{not } t} \frac{s_{v_{i,t+2}} + s_{v_{i,t+1}}}{2} (\ln \omega_{i,t+2}^v - \ln \omega_{i,t+1}^v) - \sum_{i \in t, t+1, \text{not } t+2} \frac{s_{v_{i,t+1}} + s_{vit}}{2} (\ln \omega_{i,t+1}^v - \ln \omega_{it}^v).$$

It gives the difference in growth rates between those plants absent in period t but present in $t+1$ and $t+2$ and those plants absent in period $t+2$ but present in t and $t+1$. Since the share-weights on either side of the difference do not sum to one, this term can be driven both by the prevalent growth rates of entrants and exiters and by their mass in the population.

The sum of the real productivity, reallocation, and net entry terms is equal to (22). Ignoring the net entry term, if there is no change in the growth rate of productivity at any plant, the productivity term is zero. In this case, the aggregate rate of productivity growth can only increase (decrease) if there is reallocation of the share in value added from period t to period $t+2$ towards (away from) plants with higher average productivity growth rates. Similarly, if there is no change in the shares of value added from the start (t) to the end ($t+2$), the reallocation term is zero, and the aggregate productivity growth rate can only increase if the (share-weighted) average growth rate of productivity increases. Net entry contributes a term that is positive (negative) if the share weighted sum of the growth rates of entrants exceeds (falls below) the share weighted sum of the growth rates of the exiters.

Lemma 1 provides additional motivation for this decomposition. The lemma shows that there is no reallocation effect in the aggregate productivity growth rate when markets are competitive and growth-accounting techniques are used. Thus, if a practitioner maintains perfect competition and wants to use growth-accounting techniques to infer the effect of input reallocation on welfare, he/she *must* rely on decomposing changes in the productivity growth rate.

7. The Bailey-Hulten-Campbell Index and Decomposition

In the introduction we summarized the empirical results from Section 9, which show that there is a stark divergence in the data between the BHC index and the traditional growth-accounting measure. *Fundamentally, the divergence arises because the BHC index and all of its variants*

have not been linked to welfare. Below we use a value-added formulation to demonstrate that the BHC index can be negatively correlated with welfare, can be made arbitrarily large or small by changing the units of output, uses share-weights that are incompatible with welfare measurement, and approximates the welfare change arising from entry and exit in a manner that is difficult to rationalize.

The value-added formulation defines aggregate productivity growth as the difference in share-weighted log-levels of productivity:

$$BHC_t = \sum_i s_{vit} \ln \omega_{it}^v - \sum_i s_{vi,t-1} \ln \omega_{i,t-1}^v. \quad (23)$$

BHC show that (23) can be expressed in terms of those plants that are present in both periods (continuers), and the plants that enter and exit, where the indicators denote these three types:

$$\begin{aligned} BHC_t = & \sum_i 1_{i,Continue} * (s_{vit} \ln \omega_{it}^v - s_{vi,t-1} \ln \omega_{i,t-1}^v) \\ & + \sum_i 1_{i,Enter} * s_{vit} \ln \omega_{it}^v - \sum_i 1_{i,Exit} * s_{vi,t-1} \ln \omega_{i,t-1}^v. \end{aligned} \quad (24)$$

We use continuing firms to illustrate most of our points.

BHC Reallocation Has No Theoretical Link to Welfare

For a continuing firm i , its contribution to productivity growth can be decomposed as

$$\frac{s_{vit} + s_{vi,t-1}}{2} * \Delta \ln \omega_i^v + \frac{\ln \omega_{it}^v + \ln \omega_{i,t-1}^v}{2} * \Delta s_{vi} \quad (25).$$

The first term is the growth-accounting term. It is tightly linked to welfare measurement via Lemmas 1-6 in perfectly and many imperfectly competitive settings. Indeed, those Lemmas demonstrate that this term properly measures an important part of the change in welfare arising from productivity growth in virtually every theoretical model cited in Foster et al. (2001).¹⁷ Thus, the growth-accounting component of the BHC index is well-founded in economic theory.

The problem lies with the second term in (25), which practitioners call “reallocation”. Indeed, *not one* of the models used to motivate the importance of reallocation from Foster et al. (2001) point to this term as having any precise link to welfare measurement. The consequences of this term having no theoretical tethering follow immediately.

¹⁷ This includes models of creative destruction and vintage capital like Aghion and Howitt (1992), Caballero and Hammour (1996), Campbell (1998), and many others.

BHC May Be Negatively Correlated with Welfare

It is easy to construct examples where the BHC index is negatively correlated with welfare. Let there be two firms in a competitive economy producing the same good in periods 1 and 2. One firm has higher productivity, but both firms produce because of either decreasing returns to scale or transport costs and geographically disbursed consumers. Suppose there is no change in technical efficiency for either firm from period 1 to period 2, but a distortion arises that causes the share of value added to increase at the more productive plant (e.g. a tax on the low productivity firm is enacted). *Welfare unambiguously falls* as the distortion moves the economy away from the optimal point on the production possibilities frontier. No change in technical efficiency means the first term of the BHC index (the growth-accounting term) is zero. However, the *BHC index is positive* because the reallocation term is positive; the share at the more productive plant has increased.

Sensitive to Units of Measurement

Although it has not been previously noted, it is evident from (25) that the reallocation term in the BHC index is sensitive to the units of output measurement; it is measured in *log-levels* of output. For example, if the units of output change from kilograms to grams, the BHC reallocation term increases 690% ($\ln(1000/1)$). This problem does not afflict the first term in (25) because it is defined in terms of *growth rates*. Again the BHC reallocation term is the source of the problem.¹⁸

Entry and Exit

As noted in Section 5, the contribution to the change in welfare from entrants and exiters in a growth-accounting setting is well-defined: for an entrant it is given as $\int_{t^*}^t s_{vit} d\ln\omega_{it}^v$ and for an exiter it is $\int_{t-1}^{t^*} s_{vit} d\ln\omega_{it}^v$, where t^* denotes the time of entry or exit and $s_{vit^*} = 0$.

The aggregate contribution to productivity growth from entrants and exiters in the BHC index is given in the last line of (24). The contribution of an entrant is given by $s_{vit} \ln\omega_{it}^v$, its share at t times its log-level of productivity. Similarly, the contribution of an exiter to productivity growth in the BHC index is given by $s_{v_{i,t-1}} \ln\omega_{i,t-1}^v$. If shares were constant over time, the growth-accounting measure would multiply the share times the *growth rate*, while the BHC measure multiplies the share times the *log-level*.

¹⁸ Proper price deflators for output are one solution to this problem.

Choice of Share Weight

There is an extensive discussion on the advantages and disadvantages of using labor share vs. gross-output share in the context of choosing the appropriate share weight when comparing variants of the BHC index. From a welfare perspective, this dialogue has overlooked a critical point. As Lemmas 1 and 2 show, if the desired measurement is welfare, then the ambiguity is immediately resolved. The choice of the share-weight is entirely determined by which productivity residual - gross output or value added - is used in the calculation. If the value-added residual is employed, then the share of value added is the correct weight. If i 's gross output residual is used, then i 's revenue divided by aggregate value added is the correct weight. *In neither one of these cases is the labor share or the gross output share the correct weight.*

Average Share (Tornqvist) vs. Base Period Share (Laspeyres)

This last point relates to approximation error in the “real” productivity (or growth-accounting term) in the BHC index. For a continuing plant the BHC decomposition is almost always be reported as

$$s_{v_i,t-1} \Delta \ln \omega_i^v + \ln \omega_{it}^v \Delta s_{vi} \quad (26)$$

where the first term (the “real productivity” term) is calculated using the base-period share instead of the average share. Lemmas 5 and 6 demonstrate that - without more prior information - the average share provides the best approximation. Moreover, in many settings the base-period share will provide the *worst* approximation.¹⁹ For example, if the share is monotonically increasing or decreasing between $t - 1$ and t , then the base period share is an extreme value of the set of share values visited over the period.

In results we present below (Table 4), we show that the base-period share index results in an index value that differs substantially from the average-share approximation. We interpret this difference as additional approximation error added by ignoring the period t share, the second piece of available information on the share path between $t - 1$ and t .

¹⁹ These ideas are easily formalized in a decision-theoretic framework.

8. Data and Estimation

We turn to two manufacturing censuses to explore the empirical issues that we raise. One census is from Chile's Instituto Nacional de Estadística (INE), and the second is from Colombia's Departamento Administrativo Nacional de Estadística (DANE.) The Chilean data span the period 1987 through 1996 and the Colombian data span the years 1981-1991. We focus on 3-digit industries with more than 200 observations, of which there are 23 in Chile and 26 in Colombia. Here, we provide a brief overview of these data. They have been used in numerous other productivity studies, and we refer the interested reader to those papers for a more detailed data description.²⁰

The data are unbalanced panels and cover all manufacturing plants with at least ten employees. Plants are observed annually and they include a measure of output, two types of labor, capital, and intermediate inputs. Because of the way our plant-level data are reported, we treat plants as firms, although there are probably multi-plant firms in the sample. Real value added is nominal value added adjusted by the 3-digit industry price index. Labor is the number of man-years hired for production, and plants distinguish between their blue- and white-collar workers (we include two labor types in the production function). The method for constructing the real value of capital is documented in Liu (1991) for the Chilean data, and a similar approach is adopted for the Colombian data.²¹ A data problem for the Chilean census is that approximately 3% of the plant-year observations appear to be “missing”; a plant id number is present in year $t - 1$, absent in year t , and then present again in year $t + 1$. We impute the values for these observations using $t - 1$ and $t + 1$ information (see the Appendix).

We estimate value-added production function parameters for each of the 3-digit industries and use the parameters to estimate the plant-level TFP residuals. For any industry, the production function coefficients are assumed to be constant over time and across plants, although our findings are robust to loosening this assumption. We employ three different approaches to estimating the coefficients: ordinary least squares, revenue shares, and the proxy method from Levinsohn and Petrin (2003), which includes controls to address the correlation of productivity with input choices. For the revenue shares, for each industry we use the average over plants and time of the shares.

²⁰ See Liu (1991), Liu (1993), Liu and Tybout (1996), Tybout, de Melo, and Corbo (1991), Pavcnik (2002), Levinsohn (1999), and Levinsohn and Petrin (2003).

²¹ 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.

Our intent here is *not* to compare estimators, as that has been done (again and again) elsewhere. Rather we wish to investigate whether our empirical results on measuring productivity are robust to the method of estimation.

9. Results

A necessary condition for the growth-accounting index and the BHC index to agree is that they yield similar results for continuing firms, which make up on average 95% of plant observations in the industries in our data. For these firms we compare the growth-accounting measure with the BHC index for 49 3-digit manufacturing industries from Chile and Colombia. Readers interested in the calculation of the additional terms in (16) suggested by Lemmas 3 and 4 are referred to Nishida and Petrin (2004) for a complete analysis of estimation of the full decomposition.

In an attempt to keep the analysis manageable, we start with a detailed description of results for the largest Chilean manufacturing industries. We then describe how these findings generalize. The main result is that the micro patterns observed in the largest industries in Chile are indicative of the findings for the entire 49 3-digit industries from both countries.

Table 1 reports the annual estimates of growth rates in productivity for the two measures for ISIC 311, the Food Products industry (the largest in Chile). The production function coefficients are estimated using ordinary least squares, and the productivity calculations use only plants that exist in period $t - 1$ and period t (the continuers), which in this industry account for 94.4% of plant-year observations and 96.4% of industry value added over the sample period. Column 1 is the growth in real industry value added for 1988 to 1996, column 2 is the growth-accounting measure of productivity, column 3 is the BHC productivity measure, and column 4 is the difference between these two terms, which is equal to the BHC “reallocation” term described earlier.

The growth-accounting measure averages 3.64% per annum, with standard deviation of 4.72% across the nine years. On average it accounts for slightly less than one-half of the growth rate in value added, which is consistent with the findings reported in Basu and Fernald (2002).²² The BHC index averages -2.93% per annum with a standard deviation of 13.48%. The divergence in these summary statistics arises because the two indexes themselves are widely divergent, as is evident from a comparison of columns 2 and 3. Column 4 is the difference and has a mean of -6.57%, with a standard deviation of 10.74%. Its volatility across the sample period is consistent with the general findings in the literature that “reallocation” can be large and volatile.

²² Their approach differs principally in their use of U.S. data that has been aggregated to the industry level.

For ISIC 311, table 2 compares estimates of productivity growth across different estimators for production function parameters. The top half of the table is the growth-accounting measure and the bottom half is the BHC index. For the top half, column 2 is the same as column 2 from table 1, which uses ordinary least squares to obtain production function estimates. Column 3 uses the proxy approach from Levinsohn and Petrin (2003), and column 4 uses revenue shares. While the production function estimates (not reported here) do differ somewhat across the three approaches, the productivity growth numbers using the growth-accounting index are reasonably similar across the three estimators. Only in 1995 does there seem to be substantial divergence between the revenue share estimate and the two alternatives.

For the BHC index, the signs tend to be common across the three sets of production function estimates, but the magnitudes are quite different, with the LP and OLS estimates systematically the largest in absolute value terms. The main reason for the volatility is reflected in the “reallocation” terms, which are systematically more volatile for LP and OLS relative to revenue shares.

Table 3 summarizes the relationship between the growth-accounting index and the BHC index for the eight largest industries in Chile. The industries (along with their ISIC codes) are Food Products (311), Metals (381), Textiles (321), Wood Products (331), Apparel (322), Plastics (356), Non-electric Machinery (382), and Other Chemicals (352). The index numbers use only plants that exist in $t - 1$ and t , which account for between 94% and 98% of industry value added (in this annual data, truncation due to entry and exit is not a severe problem). For each industry and each estimator of production function coefficients, the appropriate row reports the intercept, slope, and r-squared from a regression of the 9 annual growth rates using the BHC index as the independent variable and the growth-accounting index as the dependent variable. For example, the first row for ISIC 311 is the regression of column 2 on column 3 from table 1. An r-squared of 1 and an intercept of 0 and slope of 1 would indicate that the measures are identical. An r-squared of 1 would indicate that the BHC measure is a linear transformation of the growth-accounting index.

For industry 311, the r-squared's range from 0.22 to 0.47. The intercept is significantly different from zero and the slope is significantly different from one.²³ Overall, these results suggest that the BHC index is a poor proxy for the growth-accounting measure for ISIC 311. The slope terms range from 0.12 to 0.31, so the BHC index differs from the growth-accounting index by a factor of 8 using the proxy approach and 3 using revenue shares.

²³ These results do not correct for error in the parameter estimates. For this industry, with almost 10,000 observations, the parameter estimates are very precisely estimated.

These results are similar across all eight industries. R-squareds are typically low, ranging between 0.2 and 0.4. Intercept and slope coefficients are significantly different from zero and one respectively. The point estimates for the slope coefficients suggest that the BHC index differs from the growth-accounting index by several magnitudes.

The messages that come out of the results from tables 1-3 are confirmed by the other 15 manufacturing industries in Chile and the 26 manufacturing industries from Colombia. Overall, for Chile, using OLS only 4 of 23 industries had r-squareds above 0.5, using revenue shares only 5 of 23 had r-squareds above 0.5, and using the proxy approach only 1 in 23 industries had an r-squared over 0.5. Most slope coefficients across estimators and industries varied between 0.1 and 0.3.²⁴ For Colombia, using OLS only 4 of 26 industries had r-squareds above 0.5, using revenue shares 8 of 26 had r-squareds above 0.5, and using the proxy approach only 5 of 26 industries had an r-squared over 0.5. Slope coefficients across estimators and industries also varied between 0.1 and 0.3, suggesting the BHC index differs by several magnitudes from the growth-accounting index. In summary, the results demonstrate that the BHC index adds a “reallocation” term to the growth-accounting index that is large and volatile, making it a very noisy indicator of the growth-accounting index. This raises questions about the large discussion in the literature that revolves around BHC-type indexes of productivity growth.

Before turning to the decomposition that we propose, we ask whether it matters empirically if we use the base-period share (Laspeyres) or the average-share (Tornqvist) when we calculate the growth-accounting term. As we noted earlier, if the shares are changing over time at the plant-level, then using the base-period share ignores an important piece of observed information on the path of the share (the end-period share).

Table 4 compares estimated growth rates for ISIC 311. They differ substantially, as the base-period share yields an average annual rate of growth of -9.73%, while the average-share approach yields an annual growth rate of 3.64%. In particular, it seems clear from the table that the shares do vary quite a bit in a manner that is important for the approximation to the integral given in (4). The fact that the attractive results from Lemmas 5 and 6 are not available for the base-period (Laspeyres) index, and that it differs from the average-share index so dramatically suggest that researchers should regularly report the average-share approximation as one of their results. Despite its uniquely attractive features, the average share approximation is almost never reported, making it hard to compare results across studies on related questions.

²⁴ The industries that had higher r-squareds were also the industries where the slope coefficients tended to be larger.

Decomposing the Change in Productivity Growth Rates

We now turn to the decomposition of the change in productivity growth rates from Section 6. There we showed how to decompose the change in the growth rate of productivity into a term that is increasing if growth rates at the firm level are increasing (the “real productivity” term), and a term that is increasing when plants with higher average growth rates gain larger shares of value added (the “reallocation” term).

Table 5 provides this decomposition for ISIC 311 in Chile. The change in the growth rate from year to year is reported in column 2. Note that it does not exactly equal the change in the growth rate reported in column 1; the difference is due to the entry/exit plants, (those plants that do not exist in t , $t + 1$, and $t + 2$, e.g.), and it is equal to the change in the growth rate that is attributable to entry/exit. As defined in Section 6, for those plants that exist in the three periods, the change in column 2 can be decomposed into the “real productivity” term, column 3, and the “reallocation” term, column 4. The real productivity component is quite volatile, averaging -5.95% with a standard deviation of 7.33% , with both positive and negative outcomes. The reallocation term is positive, very stable, and always contributes to increases in the rate of productivity growth, even in periods when the overall growth rate falls.

Overall, the reallocation series for ISIC 311 in Chile is remarkably similar in spirit to the reallocation series from the other 48 manufacturing industries. Regardless of the estimator for production function coefficients, the industry, or the country, the annual reallocation terms are almost universally positive, and within an industry over time they vary very little, with a typical standard deviation less than 1.²⁵ Thus, these results suggest that reallocation effects - when defined in terms of changes in productivity growth rates - are very stable and almost always contribute positively to industry growth, even when the overall growth rate is falling.

10. Conclusions

In this paper we define productivity growth as the change in welfare that arises from additional output holding primary inputs constant. We show that this definition provides practitioners with a measurement useful for cost-benefit/policy analysis that is readily comparable across time, industries, countries, and empirical studies. Using traditional growth-accounting results, we show that gains may arise because of plant-level technology shocks, and, in imperfectly competitive settings, from the reallocation of inputs across plants with differing markups and/or shadow values of primary inputs.

²⁵ They do vary across industries, with most industries falling in between 3% to 8%.

The most popular index for measuring productivity growth with plant-level data uses the difference in the “first moments” of the productivity distribution, where a variety of first moments have been proposed. We show this definition adds an additional term to the growth-accounting measure, which has been called “reallocation.” Our empirical results indicate that there is a very weak relationship between the two indexes in almost every 3-digit manufacturing industry in both Chile from 1987-1996 and Colombia from 1981-1991 - 49 in total - primarily because this “reallocation” term is large and volatile. These findings are robust to different estimation approaches for plant-level productivity.

We explore the theoretical reasons for this sharp divergence, in the process uncovering a number of previously unnoticed and unattractive features of the first-moment definition: it is not linked tightly to welfare in any theoretical model of which we are aware; it is sensitive to measured units; it can report positive productivity growth when welfare has fallen; it uses labor or gross-output as share weights, neither of which is correct for welfare measurement; and its choice of the base-period versus the average share-weight leads to substantial approximation error. In the end, our results call into question the literature’s interpretation of the “reallocation” term as productivity growth. Our findings also suggest that researchers should regularly include the traditional growth-accounting measure in their results for both comparability and ease of interpretation.

References

- Abowd, J., Kramarz, F., and Margolis, D. (1999). High wage workers and high wage firms. *Econometrica*, 67(2), 251–334.
- Aghion, P., and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2), 323–351.
- Bailey, M., Hulten, C., and Campbell, D. (1992). Productivity dynamics in manufacturing plants. In *Brookings Papers on Economic Activity: Microeconomics*, Vol. 4, pp. 187–267. Brookings Institute.
- Balk, B. (2003). On the relationship between gross-output and value-added based productivity measures: The importance of the domar factor. Center for Applied Economic Research Working Paper 2003/05.
- Balk, B. (2005). Divisia price and quantity indices: 80 years after. *Statistica Neerlandica*, 59(2), 119–158.
- Basu, S., and Fernald, J. (2002). Aggregate productivity and aggregate technology. *European Economic Review*, 46, 963–991.
- Bennet, T. (1920). The theory of measurement of changes in cost of living. *Journal of the Royal Statistical Society*, 455–462.
- Bernard, A., Eaton, J., Jensen, B., and Kortum, S. (2003). Plants and productivity in international trade. *American Economic Review*, 1268–1290.
- Bernard, A., and Jensen, B. (1999). Exceptional exporter performance: Cause, effect, or both?. *Journal of International Economics*, 47, 1–25.
- Berndt, E., and Fuss, M. (1986). Productivity measurement with adjustments for variations in capacity utilization and other forms of temporary equilibrium. *Journal of Econometrics*, 33, 7–29.
- Caballero, R., and Hammour, M. (1996). On the timing and efficiency of creative destruction. *The Quarterly Journal of Economics*, 111(3), 805–852.
- Campbell, J. (1998). Entry, exit, embodied technology, and business cycles. *Review of Economic Dynamics*, 1(2), 371–408.
- Diewert, E. (1976). Exact and superlative index numbers. *Journal of Econometrics*.
- Diewert, E., and Fox, K. (2005). Measuring productivity growth. Working Paper.
- Domar, E. D. (1961). On the measurement of technological change. *The Economic Journal*, 71(284), 709–729.
- Eaton, J., Kortum, S., and Kramarz, F. (2004). An anatomy of international trade: Evidence from French firms. New York University, Department of Economics.
- 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.
- Fox, K. (2003). Problems with (dis)aggregating productivity, and another productivity paradox. University of New South Wales Discussion Paper.
- Goolsbee, A., and Petrin, A. (2004). The consumer gains from direct broadcast satellites and the competition with cable TV. *Econometrica*, 72(2), 351–382.
- Griliches, Z., and Regev, H. (1995). Firm productivity in Israeli industry: 1979–1988. *Journal of Econometrics*, 65(1), 175–203.
- Hall, R. (1990). Invariance properties of solow’s productivity residual. In Diamond, P. (Ed.), *Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow’s Birthday*. MIT Press.
- Hulten, C. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, 45(3), 511–518.

- Hulten, C. (2001). Total factor productivity: A short biography. In *New Developments in Productivity Analysis*. University of Chicago Press.
- Hulten, C. R. (1973). Divisia index numbers. *Econometrica*, 41(6), 1017–1025.
- Levinsohn, J. (1999). Employment responses to international liberalization in Chile. *Journal of International Economics*, 47, 321–244.
- 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(2), 341–372.
- Liu, L. (1991). *Entry-Exit and Productivity Changes: An Empirical Analysis of Efficiency Frontiers*. Ph.D. thesis, University of Michigan.
- Liu, L. (1993). Entry, exit, and learning in the Chilean manufacturing sector. *Journal of Development Economics*, 42, 217–242.
- Liu, 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.
- Marschak, J., and Andrews, W. (1944). Random simultaneous equations and the theory of production. *Econometrica*, 12(3-4), 143–205.
- Nishida, M., and Petrin, A. (2004). Estimating aggregate productivity growth with plant-level data when markets are not competitive. University of Chicago GSB Working Paper.
- Olley, S., and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263–1298.
- Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. *Review of Economic Studies*, 69, 245–276.
- Petrin, A. (2002). Quantifying the benefits of new products; the case of the minivan. *Journal of Political Economy*, 110, 705–729.
- Petrin, A. (2005). Use of a divisia index to aggregate productivity across firms. University of Chicago Working Paper.
- Roberts, M., and Tybout, J. (1996). *Industrial Evolution in Developing Countries*. Oxford University Press for The World Bank, New York.
- Sato, K. (1987). Ideal indexes. In *New Palgrave: A Dictionary of Economics*. New York: Stockton Press.
- Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics*, 39(3), 312–320.
- Star, S., and Hall, R. (1976). An approximate divisia index of total factor productivity. *Econometrica*, 44(2), 257–263.
- Theil, H. (1967). *Economic and Information Theory*. Rand McNally, Chicago.
- Tornqvist, L. (1936). The bank of Finland's consumption price index. *Bank of Finland Monthly Bulletin*, 1–8.
- Trivedi, P. K. (1981). Some discrete approximations to divisia integral indices. *International Economic Review*, 22(1), 71–77.
- 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 A

Proof of Lemma 1

The proof mimics Hulten. $F(\cdot)$ is differentiable and homogeneous of degree zero in (Y, X) , which implies

$$\sum_{i=1}^N \frac{\partial F}{\partial Y_i} Y_i = - \sum_{k=1}^K \frac{\partial F}{\partial X_k} X_k. \quad (27)$$

By definition $F(Y, X, t) = 0$, so total differentiation of $F(\cdot)$ yields

$$-dF = \sum_{i=1}^N \frac{\partial F}{\partial Y_i} dY_i + \sum_{k=1}^K \frac{\partial F}{\partial X_k} dX_k.$$

where dF , dY , and dX denote the instantaneous change with respect to time. Divide through by the left and right hand side of (27) and substitute in the competitive equilibrium conditions $\frac{\partial F/\partial Y_i}{\partial F/\partial Y_1} = P_i$, $i = 2, \dots, N$ and $-\frac{\partial F/\partial X_k}{\partial F/\partial Y_1} = W_k$, $k = 1, \dots, K$. Then, the shift in the social PPF holding primary inputs constant is given as the rate of change of $F(\cdot)$ divided by $\sum_{i=1}^N \frac{\partial F}{\partial Y_i} Y_i$:

$$\frac{-dF}{\sum_{i=1}^N \frac{\partial F}{\partial Y_i} Y_i} = \sum_{i=1}^N \frac{P_i Y_i}{\sum_i P_i Y_i} \frac{dY_i}{Y_i} - \sum_{k=1}^K \frac{W_k X_k}{\sum_k W_k X_k} \frac{dX_k}{X_k}.$$

This establishes the first claim.

The second claim of the Lemma asserts that this quantity is equal to $\sum_{i=1}^N \frac{P_i Q_i}{\sum_i P_i Y_i} d \ln F^i$, where $Q_i = F^i(M^i, X^i, t)$ and $d \ln F^i = d \ln \omega_i$ from (8) (we write it in this way to be consistent with the notation here). To show this, note that in equilibrium supply and demand in the product and factor markets are equated:

$$Q_i = Y_i + \sum_{j=1}^N M_i^j \quad i = 1, \dots, N$$

and

$$X_k = \sum_{j=1}^J X_k^j \quad k = 1, \dots, K.$$

Total differentiation of these equations gives

$$\frac{dQ_i}{Q_i} = \frac{dY_i}{Y_i} + \frac{\sum_{j=1}^N dM_i^j}{Q_i} = \frac{P_i Y_i}{P_i Q_i} \frac{dY_i}{Y_i} + \sum_{j=1}^N \frac{P_i M_i^j}{P_i Q_i} \frac{dM_i^j}{M_i^j} \quad i = 1, \dots, N \quad (28)$$

and

$$\frac{dX_k}{X_k} = \sum_{i=1}^N \frac{W_k X_k^i}{W_k X_k} \frac{dX_k^i}{X_k^i} \quad k = 1, \dots, K. \quad (29)$$

Necessary conditions for the equilibrium are: $\frac{\partial Q_i}{\partial M_j^i} = \frac{P_i}{P_j}$ and $\frac{\partial Q_i}{\partial X_k^i} = \frac{W_k}{P_i}$, $i, j = 1, \dots, N$ and $k = 1, \dots, K$. Substituting these equalities into the logarithmic derivative of $Q_i = F^i(M^i, X^i, t)$ and dividing through by Q_i yields

$$\frac{dQ_i}{Q_i} = \sum_{j=1}^N \frac{P_j M_j^i}{P_i Q_i} \frac{dM_j^i}{M_j^i} + \sum_{k=1}^K \frac{W_k X_k^i}{P_i Q_i} \frac{dX_k^i}{X_k^i} + d\ln F^i \quad i = 1, \dots, N, \quad (30)$$

where $d\ln F^i$ is the well-known Solow residual. Solving (28) for $\frac{P_i Y_i}{P_i Q_i} \frac{dY_i}{Y_i}$ and plugging in (30) yields

$$\frac{P_i Y_i}{P_i Q_i} \frac{dY_i}{Y_i} = \sum_{j=1}^N \frac{P_j M_j^i}{P_i Q_i} \frac{dM_j^i}{M_j^i} + \sum_{k=1}^K \frac{W_k X_k^i}{P_i Q_i} \frac{dX_k^i}{X_k^i} + d\ln F^i - \sum_{j=1}^N \frac{P_j M_j^i}{P_i Q_i} \frac{dM_j^i}{M_j^i} \quad i = 1, \dots, N \quad (31)$$

Multiplying (31) through by $\frac{P_i Q_i}{\sum_{i=1}^N P_i Y_i}$ and noting from the accounting identity $\sum_{i=1}^N P_i Y_i = \sum_{k=1}^K W_k X_k$ we have

$$\frac{P_i Y_i}{\sum_i P_i Y_i} \frac{dY_i}{Y_i} = \sum_{j=1}^N \frac{P_j M_j^i}{\sum_i P_i Y_i} \frac{dM_j^i}{M_j^i} + \sum_{k=1}^K \frac{W_k X_k^i}{\sum_k W_k X_k} \frac{dX_k^i}{X_k^i} + \frac{P_i Q_i}{\sum_i P_i Y_i} d\ln F^i - \sum_{j=1}^N \frac{P_j M_j^i}{\sum_i P_i Y_i} \frac{dM_j^i}{M_j^i}. \quad (32)$$

Aggregating across i gives

$$\sum_{i=1}^N \frac{P_i Y_i}{\sum_i P_i Y_i} \frac{dY_i}{Y_i} = \sum_{i=1}^N \sum_{j=1}^N \frac{P_j M_j^i}{\sum_i P_i Y_i} \frac{dM_j^i}{M_j^i} + \sum_{i=1}^N \sum_{k=1}^K \frac{W_k X_k^i}{\sum_k W_k X_k} \frac{dX_k^i}{X_k^i} + \sum_{i=1}^N \frac{P_i Q_i}{\sum_i P_i Y_i} d\ln F^i - \sum_{i=1}^N \sum_{j=1}^N \frac{P_j M_j^i}{\sum_i P_i Y_i} \frac{dM_j^i}{M_j^i}. \quad (33)$$

The intermediate inputs cancel out because

$$\sum_{i=1}^N \sum_{j=1}^N P_j dM_j^i = \sum_{i=1}^N \sum_{j=1}^N P_i dM_i^j.$$

Rearranging (33) then yields

$$\sum_{i=1}^N \frac{P_i Q_i}{\sum_i P_i Y_i} d\ln F^i = \sum_{i=1}^N \frac{P_i Y_i}{\sum_i P_i Y_i} \frac{dY_i}{Y_i} - \sum_{k=1}^K \frac{W_k X_k}{\sum_k W_k X_k} \frac{dX_k}{X_k}$$

This establishes the second claim. #

Proof of Lemma 2

Let $s_{M^i} = \frac{\sum_{j=1}^N P_j M_j^i}{P_i Q_i}$, the payments to intermediates divided by gross output. The residual from the value-added production function is given by $d\ln F^{iV} \equiv \frac{d\ln F^i}{(1-s_{M^i})}$; the change in plant-level efficiency $d\ln F^i$ is grossed up by $\frac{1}{(1-s_{M^i})}$ to account for the role of intermediates on final demand.

Because $P_i Y_i = P_i Q_i - \sum_j P_j M_j^i$, it follows directly that

$$\sum_{i=1}^N s_{v_i} d\ln F^{iV} = \sum_{i=1}^N \frac{P_i Y_i}{\sum_{i=1}^N P_i Y_i (1-s_{M^i})} d\ln F^i = \sum_{i=1}^N \frac{P_i Y_i}{\sum_{i=1}^N P_i Y_i} \frac{d\ln F^i}{\frac{P_i Q_i - \sum_j P_j M_j^i}{P_i Q_i}} = \sum_{i=1}^N \frac{P_i Q_i}{\sum_i P_i Y_i} d\ln F^i.$$

#

Lemma 3: A Simple Example

The simple two-period general equilibrium model given in Basu and Fernald (2002) illustrates how the envelope theorem links productivity to welfare. Consider a competitive economy with optimizing agents: a representative consumer and firm, prices taken as given, a constant returns production function $Q_t = \omega_t F(K_t, L_t)$, and an initial endowment of capital K_1^* , and two independent productivity shocks (ω_1, ω_2) . How much does welfare change in response to a shock $d\omega_1$?

This problem is isomorphic to the social planner's problem, who maximizes welfare:

$$\text{Max}_{C_1, C_2, L_1, L_2} V(K_1^*, \omega_1, \omega_2) = E_1[U(C_1, 1 - L_1) + \beta U(C_2, 1 - L_2)]$$

subject to the budget constraints

$$C_1 = Q_1 - I_1 = \omega_1 F(K_1^*, L_1) - (K_2 - (1 - \delta)K_1^*)$$

and

$$C_2 = \omega_2 F(K_2, L_2) + (1 - \delta)K_2,$$

where I_1 denotes investment (savings) in period 1. Let $(C_1^*, C_2^*, L_1^*, L_2^*)$ denote the solution and $V^* = V(C_1^*, C_2^*, L_1^*, L_2^*; \omega_1)$.

The envelope condition says that the total change in maximal utility dV^* due to a small change in parameter $d\omega_1$ is equal to the partial effect evaluated at the initial optimum:

$$\frac{dV^*}{d\omega_1} = \frac{\partial V(C_1^*, C_2^*, L_1^*, L_2^*; \omega_1)}{\partial \omega_1}.$$

Using the Lagrange formulation for the constrained problem, it is straightforward to show that the marginal utility of wealth is equal to the partial derivative of utility with respect to C_1 , when evaluated at the optimum:

$$\frac{dV^*}{dQ_1} = \frac{\partial U(C_1^*, L_1^*)}{\partial C_1} = \frac{\partial U^*}{\partial C_1}.$$

Inserting the above formulation for C_1 and C_2 into $V(\cdot)$ yields the unconstrained problem, and differentiating with respect to ω_1 gives

$$\frac{\partial V^*}{\partial \omega_1} = \frac{\partial U^*}{\partial C_1} * F(K_1^*, L_1^*).$$

The total change in utility is then:

$$dV^* = \frac{\partial U^*}{\partial C_1} * F(K_1^*, L_1^*) * d\omega_1 = \frac{\partial U^*}{\partial C_1} * Q_1^* * d\ln\omega_1.$$

The additional units of output available is given by the growth rate of productivity times the level of output, and the marginal utility of wealth maps this additional output into utility. The conclusion that the change in welfare is proportional to productivity growth follows directly.

Proof of Lemma 4

Let Z_j denote a generic input with P_j^Z its price, and let $Z = (Z_1, \dots, Z_J)$ equal the full vector of inputs. Firm i minimizes cost while achieving the target output Q^* :

$$\text{Min}_{Z_1, Z_2, \dots, Z_J} \sum_j P_j^Z Z_j$$

such that

$$F^i(Z) \geq Q^*.$$

The Lagrangian is

$$L = \sum_j P_j^Z Z_j - \lambda[F^i(Z) - Q^*]$$

with first order conditions (FOC):

$$\frac{\partial L}{\partial Z_j} = P_j^Z - \lambda \frac{\partial F^i(Z)}{\partial Z_j}$$

for $j = 1, \dots, J$, and

$$\frac{\partial L}{\partial Q^*} = \lambda.$$

With $\lambda = MC_i$, the input FOCs can be rewritten as

$$\frac{P_i}{P_i} \frac{P_j^Z}{MC_i} \frac{Z_j}{F^i(Z)} = \frac{\partial F^i(Z)}{\partial Z_j} \frac{Z_j}{F(Z)}.$$

Reexpressed we have

$$\mu_i s_j = \epsilon_{Z_j},$$

for $j = 1, \dots, J$, where optimizing firms equate - for each input - the product of the markup and revenue share with the elasticity of output with respect to the input. With three inputs labor, capital, and materials production can be written as

$$d \ln Q_i = \sum_i \mu_i (s_{L^i} d \ln L^i + s_{K^i} d \ln K^i + s_{M^i} d \ln M^i) + d \ln \omega_i, \quad (34)$$

a generalization of Solow's formulation to the case of imperfect competition.

(13) can be reexpressed as

$$d \ln Y_i = d \ln Q_i - \frac{s_{M^i}}{1 - s_{M^i}} (d \ln M_i - d \ln Q_i).$$

From (14), all that remains to be shown is

$$d \ln Q_i = \mu_i^v d X_i^v + \mu_i^v \frac{s_{M^i}}{1 - s_{M^i}} (d \ln M^i - d \ln Q_i) + \frac{d \ln \omega_i}{1 - \mu_i s_{M^i}}.$$

If we write $d\ln Q_i$ as

$$d\ln Q_i = \frac{d\ln Q_i}{1 - \mu_i s_{M^i}} - \frac{\mu_i s_{M^i}}{1 - \mu_i s_{M^i}} d\ln Q_i,$$

and replace the numerator of first term on the right of the equality with (34), the claim is established.

(15) follows directly from the first result, and the equality of (16) and (12) is easily shown using

$$1 - s_{M^i} = \frac{P_i Q_i}{P_i Y_i}$$

and

$$P_i Y_i * d\ln \omega_i^v = P_i dY_i - P_{L^i} L^i d\ln L^i - P_{K^i} K^i d\ln K^i.$$

#

Appendix B

Imputing Missing Values

Approximately 3% of the plant-year observations in Chile are “missing” according to the following definition: a plant id number is present in year $t - 1$, absent in year t , and then present again in year $t + 1$. We impute the values for these observations using $t - 1$ and $t + 1$ information and the structure of the estimated production function. We use the simple average of the $t - 1$ and $t + 1$ (log) productivity estimates for the period t productivity estimate. Similarly, we use the simple average of the $t - 1$ and $t + 1$ (log) input index estimates, where the weights in the index are the estimated production function parameters. All of our findings are robust to dropping these observations.

TABLE 1
 Comparison of the Growth-Accounting
 Index with the BHC Productivity Index
 Ordinary Least Squares Estimates, ISIC 311, Chile
 Rate of Growth in:

Year	Value Added	Growth-Accting Index $\sum_i \frac{(s_{it} + s_{i,t-1})}{2} \ln\left(\frac{\omega_{it}}{\omega_{i,t-1}}\right)$	BHC Index $\sum_i s_{it} \ln \omega_{it} - \sum_i s_{i,t-1} \ln \omega_{i,t-1}$	Difference (BHC "Reallocation" Term) $\sum_i \frac{(\ln \omega_{it} + \ln \omega_{i,t-1})}{2} * \Delta s_i$
1988	11.75	-3.12	-14.96	-11.84
1989	7.36	3.64	-7.45	-11.10
1990	4.59	-1.10	-10.52	-9.43
1991	13.82	7.36	-13.28	-20.63
1992	14.67	6.09	-8.11	-14.20
1993	8.98	5.09	-0.19	-5.28
1994	8.20	2.95	7.09	4.13
1995	7.06	-0.74	-7.24	-6.50
1996	-1.30	12.56	28.25	15.69
Average	8.35	3.64	-2.93	-6.57
Std. Dev.	4.89	4.72	13.48	10.74

The last column is the discrepancy between the BHC index and the growth-accounting index, which is equal to a reallocation-like term given by $\sum_i \overline{\ln \omega_i} * \Delta s_i$. The comparison is done on firms that exist in period t and $t - 1$, which account for 94.4% of the plant-year observations and 96.4% of industry value added. See text for details.

TABLE 2
 Comparison of Productivity Indexes Across
 OLS, Levinsohn-Petrin, and Revenue Share Productivity Estimates
 ISIC 311, Chile

Year	Value Added	Growth-Accting Index		Revenue Shares
		OLS	Levinsohn- Petrin	
1988	11.75	-3.12	-0.04	0.77
1989	7.36	3.64	4.39	4.86
1990	4.59	-1.10	-1.26	-1.91
1991	13.82	7.36	7.68	8.98
1992	14.67	6.09	6.82	9.69
1993	8.98	5.09	5.87	4.41
1994	8.20	2.95	4.45	2.84
1995	7.06	0.74	1.41	-5.22
1996	-1.30	12.56	11.07	6.80
		BHC Index		
1988	11.75	-14.96	-19.16	-5.04
1989	7.36	-7.45	-11.94	-4.40
1990	4.59	-10.52	-20.43	-2.21
1991	13.82	-13.28	-24.42	-4.09
1992	14.67	-8.11	-15.54	-0.01
1993	8.98	-0.19	5.16	0.09
1994	8.20	7.09	4.81	4.33
1995	7.06	-7.24	-9.09	-10.40
1996	-1.30	28.25	39.66	1.07

Growth rates in productivity compared across methods used to estimate production function parameters. Comparison is done on firms that exist in period t and $t - 1$.

TABLE 3
 Regression of Growth-Accounting Productivity Index on BHC Productivity Index
 Chilean Manufacturing, Industry by Industry, 1988-1996 (9 observations)

Industry Code	Coefficient Estimator	Intercept (Std. Err.)	Slope (Std. Err.)	R-squared
311	OLS	.043 (.012)	.247 (.099)	0.47
	LP	.051 (.011)	.122 (.058)	0.38
	Rev.	.036	.314	0.22
	Shares	(.015)	(.223)	
381	OLS	.031 (.023)	.228 (.091)	0.47
	LP	.028 (.024)	.230 (.096)	0.44
	Rev.	.008	.373	0.62
	Shares	(.021)	(.109)	
321	OLS	.010 (.023)	.177 (.129)	0.21
	LP	.017 (.020)	.092 (.085)	0.14
	Rev.	-.009	.503	0.38
	Shares	(.026)	(.026)	
331	OLS	.017 (.046)	.207 (.132)	0.25
	LP	.018 (.047)	.146 (.108)	0.20
	Rev.	.043	.501	0.57
	Shares	(.045)	(.163)	

Row one for ISIC 311 is the regression of column 2 on column 3 from table 1 (for example). An r-squared of 1 and an intercept of 0 and slope of 1 would indicate that the measures are identical. An r-squared of 1 would indicate that the BHC measure is a linear transformation of the growth-accounting index.

TABLE 3 (continued)
 Regression of Growth-Accounting Index on BHC Productivity Index
 Chilean Manufacturing, Industry by Industry, 1988-1996 (9 observations)

Industry Code	Coefficient Estimator	Intercept (Std. Err.)	Slope (Std. Err.)	R-squared
322	OLS	-0.008 (.031)	.223 (.154)	0.22
	LP	-0.008 (.033)	.214 (.156)	0.21
	Rev. Shares	-0.008 (.033)	.482 (.201)	0.44
356	OLS	-0.047 (.029)	.024 (.061)	0.02
	LP	-0.043 (.031)	.028 (.066)	0.02
	Rev. Shares	-0.22 (.042)	.165 (.143)	0.15
382	OLS	.093 (.036)	.068 (.063)	0.14
	LP	.093 (.037)	.058 (.062)	0.11
	Rev. Shares	.072 (.047)	.180 (.137)	0.19
352	OLS	.019 (.031)	.274 (.168)	0.27
	LP	.039 (.031)	.156 (.128)	0.17
	Rev. Shares	.003 (.024)	.416 (.155)	0.50

Row one for ISIC 311 is the regression of column 2 on column 3 from table 1 (for example). An r-squared of 1 and an intercept of 0 and slope of 1 would indicate that the measures are identical. An r-squared of 1 would indicate that the BHC measure is a linear transformation of the growth-accounting index.

TABLE 4
 Comparison of Base-Period Share to Average Share
 In Approximating the Shift in the Growth-Accounting Index
 Ordinary Least Squares Estimates, ISIC 311, Chile

Year	Rate of Growth in:		
	Value Added	Average Share (Tornquist): $\sum_i \frac{(s_{it} + s_{i,t-1})}{2} \ln\left(\frac{\omega_{it}}{\omega_{i,t-1}}\right)$	Base Period Share (BHC) (Laspeyres): $\sum_i s_{i,t-1} \ln\left(\frac{\omega_{it}}{\omega_{i,t-1}}\right)$
	1988	11.75	-3.12
1989	7.36	3.64	-11.17
1990	4.59	-1.10	-16.12
1991	13.82	7.36	-4.03
1992	14.67	6.09	-7.19
1993	8.98	5.09	-8.61
1994	8.20	2.95	-10.12
1995	7.06	-0.74	-8.48
1996	-1.30	12.56	0.02
Average	8.35	3.64	-9.73
Std. Dev.	4.89	4.72	6.41

The comparison is done on firms that exist in period t and $t - 1$, which account for 94.4% of the plant-year observations and 96.4% of industry value added. See text for details.

TABLE 5
Petrin and Levinsohn Decomposition of
the Change in Rate of Productivity Growth
“Real Productivity” vs. “Reallocation”

Ordinary Least Squares Estimates, ISIC 311, Chile

Year	Productivity Growth Rate	Change in Growth Rate	Real Productivity Component	Reallocation Component
1988	-3.12	—	—	—
1989	3.64	7.40	-1.79	9.18
1990	-1.10	-5.25	-14.83	9.59
1991	7.36	10.86	2.35	8.51
1992	6.09	-2.30	-10.64	8.34
1993	5.09	-0.66	-9.17	8.51
1994	2.95	-1.12	-9.72	8.60
1995	-0.74	-2.44	-10.08	7.64
1996	12.56	13.61	6.23	7.38
Average	3.64	2.51	-5.95	8.46
Std. Dev.	4.72	7.05	7.33	0.72

The first column is the rate of productivity growth from $t - 1$ to t estimated using the growth-accounting index. The second column is the change in this growth rate for firms that exist in $t - 2$, $t - 1$, and t . The discrepancy between the change in column 1 and the level in column 2 is due to firms that do not exist in all three periods (the net entry term). The third and fourth column decompose column 2. See text for details.