Measuring Aggregate Productivity Growth Using Plant-level Data

by

Amil Petrin
University of Minnesota, Twin Cities
National Bureau of Economic Research
and
James Levinsohn
Yale University
National Bureau of Economic Research

March 24, 2011

Abstract. We define aggregate productivity growth as the change in aggregate final demand minus the change in the aggregate cost of primary inputs. We show how to aggregate plant-level data to this measure and how to use plant-level data to decompose our measure into technical efficiency and reallocation components. This requires us to confront the "non-neoclassical" features that impact plant-level data including plant-level heterogeneity, the entry and exit of goods, adjustment costs, fixed and sunk costs, and market power. We compare our measure of aggregate productivity growth to several competing variants that are based only on a single plant-level factor of technical efficiency. We show that theory suggests our measure may differ substantially from these measures of aggregate productivity growth. We illustrate this using panel data from manufacturing industries in Chile. We find that our measure does differ substantially from other widely used measures with especially marked differences in the fraction of productivity growth attributed to reallocation.

Measuring Aggregate Productivity Growth Using Plant-level Data

Amil Petrin
University of Minnesota, Twin Cities
National Bureau of Economic Research
and
James Levinsohn
Yale University
National Bureau of Economic Research

March 24, 2011

1. Introduction

We define aggregate productivity growth (APG) as the change in the value of aggregate final demand minus the change in total expenditures on labor and capital. Our contribution is to develop an accounting framework that aggregates plant-level measurements of inputs and outputs to this quantity. We do so while accounting for non-neoclassical features of the plant-level economic environment such as differences in plant-level production technologies and productivity levels, the entry and exit of goods, adjustment costs for inputs or outputs, fixed and sunk costs, and market power. More generally, any economic phenomenon affecting aggregate final demand or the total expenditures on labor or capital inputs must - by definition - be accounted for in our framework.

Our results extend Hulten (1978) and Basu and Fernald (2002), who link the production side to changes in aggregate final demand using the National Income Identity: aggregate value-added equals aggregate final demand because intermediate input usage cancels out (in the aggregate). Hulten (1978) shows how to aggregate plant-level technical efficiency shocks to the change in final demand in the neoclassical setting, where relative prices are aligned with marginal rates of substitution and marginal rates of transformation. Basu and Fernald (2002) use the structure from Hall (1990) to extend Hulten (1978) to allow for markups. In contrast to the Hulten (1978) setting, they show that the existence of markups means growth can occur if inputs are reallocated from low markup to high markup firms.

One drawback of these frameworks is that they are not able to accommodate departures from the neoclassical setting that lead to kinks or jumps in APG. This rules out the entry and exit of new goods, which are known to play an important role in growth (see Trajtenberg (1989), Boskin Commission (1996), Petrin (2002), and Goolsbee and Petrin (2004)). Kinks and jumps in APG arise in s-S type models for labor or capital that lead to lumpy adjustment, as in Bentolila and Bertola (1990), Caballero and Engel (1993)

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."

or Caballero and Engel (1999). They are also present in the creative destruction models of Aghion and Howitt (1992) or Lentz and Mortensen (2008), and in vintage capital models with non-exponential rates of depreciation (Benhabib and Rustichini (1991)). In this paper we develop a definition of APG that is in the spirit of the original Solow (1957) residual and that does not rule out these and other important "non-differentiable"features of the micro-level.

Aggregate final demand can increase without an increase in input use if a plant becomes more technically efficient– that is, if the plant produces more output with the same inputs. This improvement might be the invention of a new method of production or by learning to imitate other better-performing plants. We show that the increase in APG due to a technical efficiency gain is equal to the plant's additional output multiplied by its price (or value). In growth rates, APG changes by the sum of weighted plant-level changes in technical efficiency growth, with the weight equal to the ratio of plant-level revenue to aggregate final demand (the Domar (1961) weight).

Final demand can also increase without more input use if non-neoclassical factors lead to differences across plants in the value of the marginal output for some input. In our setup, final demand increases if an input is *reallocated* from a lower marginal value activity to a higher marginal value one. More generally, when APG is defined in terms of changes in final demand, we show that any unit increase in any input increases APG by that input's concurrent *value of marginal product-input cost* gap.

Since our definition is based on plant-level measurements that allow for entry and exit and add up exactly to APG, we are able to characterize the source of every plant's contribution to APG. We can thus group different types of firms– e.g. entrants and exiters or exporters and non-exporters - and directly compare their contributions to growth via reallocation or technical efficiency. In addition to shedding light on the key mechanisms of economic growth, our decomposition can be used to evaluate theoretical models of growth that have testable predictions based on changes in technical efficiency or reallocation.[1]

There is a large literature on estimating plant-level productivity, and we are mostly agnostic in this paper as to the competing approaches. Rather, we use several different estimators and then aggregate these plant-level estimates to APG. Just as there are competing approaches to estimating plant-level productivity, there are several alternative definitions of APG. These include Bailey, Hulten, and Campbell (1992) (BHC) and its derivatives (e.g. Griliches and Regev (1995b), Olley and Pakes (1996), and Foster, Haltiwanger, and Krizan (2001)). None of these alternatives is directly linked to changes in aggregate final demand. Instead they are defined as output- or input-share weighted changes in the distribution of plant-level technical efficiencies. When so defined, reallocation growth equals the change in share multiplied by the technical efficiency level. For many questions this definition is problematic because increases in "reallocation"can be associated with *decreases* in final demand. This is because technical efficiency itself is a production concept and has nothing to do with the value of additional output associated with technical efficiency gains. For

[1] On growth from reallocation see also Hopenhayn and Rogerson (1993), Caballero and Hammour (1994), Mortensen and Pissarides (1994), Campbell (1998), Melitz (2003), and Restuccia and Rogerson (2007).

example, an economy may be endowed with an extremely efficient production technology for widgets, but that does not mean large increases in final demand will result if resources are reallocated towards widget production. Indeed if resources are reallocated from more highly valued activities, aggregate final demand will fall.

In the next section, we develop our definition of APG. In section 3, we show how to decompose our measure into components capturing technical efficiency, reallocation, and fixed costs. Section 4 discusses how our measure of productivity growth is impacted by particular features of the data. In section 5, we compare the decomposition of our measure of APG with the measure currently used in much of the literature and we show, theoretically, why the two measures may yield substantively different conclusions. To apply our measure to data, one must move from the continuous time set-up used in our model to a discrete time framework. This is done in section 6, while section 7 presents the data and estimation. Results are collected in section 8, and section 9 concludes.

2. Aggregate Productivity Growth

We consider a continuous-time setting that has N potential goods indexed by i. For transparency, we suppress the time index when possible and we assume each plant in the economy produces one good (the extension to a multi-product setting is possible).[2] Every product i may have a different technology, and we express the production function for good i as

$$Q^i(X_i, M_i, \omega_i).$$

The collection of primary inputs K used in production at plant i is denoted $X_i = (X_{i1}, \ldots, X_{iK})$, with X_{ik} the amount of the kth primary input used. Any product i may potentially be used as an intermediate input in the production of itself or another product, with M_{ij} denoting the amount of j's output used as an input in the production of i, and the entire vector of intermediate inputs used in production at plant i denoted $M_i = (M_{i1}, \ldots, M_{iN})$. ω_i denotes the technical efficiency at plant i.

We use F_i to denote the sum of all fixed and sunk costs incurred at plant i at any particular time. Given the wide variety of fixed and sunk costs that arise in the theoretical literature, we formulate F_i as

$$F_i = \sum_k F_{ik} + \sum_j F_{ij} + F_{i0}.$$

These terms play an important role in the economic growth literature and can lead to kinks and jumps in APG. For example, F_{ik} or F_{ij} may reflect fixed/sunk costs associated with using any primary or intermediate input respectively. In creative destruction/vintage capital models F_{ii} would reflect the sunk costs associated with developing the new vintages or new products, and they might be one-time entry/development costs or

[2] We abstract from measurement issues and we treat inventories, investment, and exports and imports in Section 4.

continually paid flow costs like ongoing research and development.[3] In search cost and hiring and firing cost models, F_{ik} would reflect the adjustment costs of finding and training a new employee, and the severance payments and potential legal costs associated with firing employees.[4] Finally, we let F_{i0} capture all costs of production that do not show up elsewhere in this accounting framework.

We follow the theoretical literature, normalizing the sum of all fixed and sunk costs to their equivalent of forgone output and deducting them directly from Q^i:

$$Q_i = Q^i(X_i, M_i, \omega_i) - F_i.$$

We collect primary and intermediate inputs and productivity shocks for firm i in $Z_i = (X_i, M_i, \omega_i)$. For the entire economy we write $Z = (Z_1, Z_2, \ldots, Z_N)$. Given Z, output quantities are determined by the production technologies, and we write $Q = (Q_1(Z_1), \ldots, Q_N(Z_N))$. We assume prices are uniquely determined by Q, given as $P = (P_1(Q_1), \ldots, P_N(Q_N))$, and similarly for primary input costs $W = (W_1(Z), \ldots, W_K(Z)))$. We assume fixed and sunk costs for all i are deterministic given Z and its past values, and we collect these fixed costs in the vector $F = (F_1, \ldots, F_N)$.

Given Q^i and F_i, the total amount of output from plant i that goes to final demand Y_i is then

$$Y_i = Q_i - \sum_j M_{ji},$$

where $\sum_j M_{ji}$ is the total amount of i's output that serves as intermediate input within the plant and at other plants. With $dY_i = dQ_i - \sum_j dM_{ji}$, the change in aggregate final demand is

$$\sum_{i=1}^{N} P_i dY_i,$$

the value of additional output excluding that output used as intermediate inputs. APG, defined as the change in aggregate final demand minus the change in aggregate expenditures on labor and capital, is then equal to:

$$APG \equiv \sum_{i=1}^{N} P_i dY_i - \sum_{i=1}^{N} \sum_k W_{ik} dX_{ik}, \tag{1}$$

with changes in primary input use reflected in $\sum_{i=1}^{N} \sum_k W_{ik} dX_{ik}$.Entry and exit is directly encompassed in this continuous time setup (see the Appendix for discussion and examples). 1from time 0 to time 1 is given by the integral:

$$APG[0,1] \equiv \int_0^1 (\sum_i P_i(t) dY_i(t) - \sum_i \sum_k W_{ik}(t) dX_{ik}(t)) dt. \tag{2}$$

[3] See e.g. Ericson and Pakes (1995), Bresnahan and Reiss (1991), Caballero and Hammour (1994), Aghion and Howitt (1994).

[4] See Mortensen and Pissarides (1994), Caballero and Hammour (1996), or Petrin and Sivadasan (2010).

While our framework does not hinge on APG equalling the change in welfare, we *can* use Basu and Fernald (2002) to illustrate conditions under which our APG measure does link to welfare.[5] We simplify their proof to illustrate the idea. Assuming continuous demands for N consumption goods $Y = (Y_1, \ldots, Y_N)$ and abstracting from investment and labor supply decisions, the consumer maximizes utility $U(Y)$ subject to income I:

$$Max\ U(Y)\ \ s.t. \sum_i P_i Y_i \leq I.$$

The Lagrangian is given as $[MaxU(Y) + \lambda(I - \sum_i P_i Y_i)]$, and assuming an interior solution exists, the vector of demands Y^* satisfies the first-order conditions

$$\frac{1}{\lambda}\frac{\partial U(Y^*)}{\partial Y_i} = P_i \ \ i = 1, \ldots N.$$

It follows that

$$\sum_i P_i dY_i = \frac{1}{\lambda} * \sum_i \frac{\partial U(Y^*)}{\partial Y_i} dY_i,$$

and the change in final demand is proportional to the change in utility with the constant of proportionality equal to $\frac{1}{\lambda}$, the inverse of the marginal utility of income.

In plant-level data sets, we observe plant-level value-added, but we do not observe the amount of a plant's output that ultimately goes to final demand. Calculation of APG is then not possible using (1) because of the unavailability of dY_i for every plant i. The usefulness of plant-level data arises from the National Accounting Identity, which shows that aggregate final demand is equal to aggregate value-added:

$$\sum_i P_i Y_i = \sum_i VA_i \tag{3}$$

with value-added

$$VA_i = P_i Q_i - \sum_j P_j M_{ij},$$

and

$$dVA_i = P_i dQ_i - \sum_j P_j dM_{ij}.$$

Given (3) we can replace the first term in (1) and calculate aggregate productivity growth as

$$APG = \sum_i dVA_i - \sum_i \sum_k W_{ik} dX_{ik}. \tag{4}$$

[5] Our definition of APG is closest to Basu and Fernald (2002). By restricting themselves to a framework which is differentiable they can show their definition provides a first-order approximation to welfare. The drawback is that they cannot allow for the entry and exit of new goods nor can they explicitly incorporate fixed and sunk costs that lead to kinks and jumps for APG. A second difference is they assume common input prices across firms, and that these input prices exactly equal the shadow value of the input. This leads them to define the change in aggregate expenditures in terms of changes in aggregated capital and aggregated labor, as opposed to the micro-level changes themselves as in (1).

3. Decomposing APG

It is straightforward to decompose APG into a component related to the aggregate of plant-level technical efficiencies and the aggregate of several plant-level input reallocation terms, one for each input. Lemma 1 provides the decomposition for APG by differentiating Q_i, where we normalize $\frac{\partial Q_i}{\partial \omega} = 1$.

Lemma 1

Assume $Q_i(\cdot)$ is once differentiable for all i. Let

$$TE \equiv \sum_i P_i d\omega_i, \tag{5}$$

$$RE \equiv \sum_i \sum_k (P_i \frac{\partial Q_i}{\partial X_k} - W_{ik}) dX_{ik} + \sum_i \sum_j (P_i \frac{\partial Q_i}{\partial M_j} - P_j) dM_{ij}, \tag{6}$$

and

$$F \equiv -\sum_i P_i dF_i \qquad (7).$$

Then

$$APG = TE + RE + F. \tag{8}$$

See the Appendix for proof. The Technical Efficiency (TE) term is the contribution to aggregate productivity growth coming from plants generating more output holding inputs constant. The Fixed Costs (F) term deducts all incurred fixed and sunk costs and the reallocation term (RE) relates changes in input allocation across plants to changes in final demand. The equality holds when the sum is taken over all plants in the economy. It also holds for any subset of plants, where one recovers for that subset their contribution to aggregate technical efficiency growth (5) and aggregate reallocation (6). This is useful for evaluating the different sources of growth arising from different subsets of plants (e.g. recent entrants/exiters or exporters/non-exporters).

The reallocation terms are based on the value of the marginal products (VMP) for every input, given generically for any input X_k at firm i as:

$$VMP_{ik} \equiv P_i \frac{\partial Q_i}{\partial X_k}.$$

The reallocation terms include a VMP term and an input cost term for each plant and every primary and intermediate input. Using labor as an example, assuming common wages, reallocation of a unit of labor from j to i would lead $dL_i = 1$ and $dL_j = -1$, and would thus increase the value of output by

$$P_i \frac{\partial Q_i}{\partial L} - P_j \frac{\partial Q_j}{\partial L}$$

while holding total labor input constant. Thus, if an input moves from a low marginal value activity to a higher one, aggregate final demand increases without any increase in technical efficiency or aggregate input use. The marginal unit of output from plants with bigger gaps between price and marginal cost add more to APG as their output increases, relatively speaking, and aggregate reallocation increases when inputs move from plants with smaller gaps to those with bigger gaps.

Markups, taxes, and adjustment costs all create gaps. The markup is by definition the gap between price and marginal cost, and APG increases when output increases at any plant with price above marginal cost. APG also increases when resources shift from plants with low markups to plants with high markups.A tax of τ on a good reduces the marginal revenue of the kth input from $P_i \frac{\partial Q}{\partial X_k}$ to $\frac{1}{1+\tau} P_i \frac{\partial Q}{\partial X_k}$, so plants produce at a level of output where $P_i \frac{\partial Q}{\partial X_k} > W_{ik}$. Costs of adjustment for inputs lead to s-S type models, where there are ranges of demand or technical efficiency shocks such that the plant does not necessarily adjust inputs in every period, and when they do adjust, they do not adjust to equate the value of the marginal product with the input price. [6] These are just some examples of economic phenomena that lead to a role for reallocation in changes in aggregate final demand, and the RE term provides the basis for the calculation of the impact of policies on gaps and their relationship to movements in inputs (see Section 4).

The growth rate formulation also follows from Lemma 1. Output elasticities with respect to input m are denoted as ε_{im} and revenue shares are denoted $s_{ik} = \frac{W_{ik} X_{ik}}{P_i Q_i}$.

Corollary 1

In growth rates (8) is given as

$$APG_G = \sum_i D_i dln\omega_i + \sum_i D_i \sum_k (\varepsilon_{ik} - s_{ik}) dlnX_{ik} + \sum_i D_i \sum_j (\varepsilon_{ij} - s_{ij}) dlnM_{ij} - \sum_i D_i dlnF_i, \qquad (9)$$

where the Domar weight is $D_i = \frac{P_i Q_i}{\sum_{i=1}^N P_i Y_i} = \frac{P_i Q_i}{\sum_{i=1}^N VA_i}$, ε_{ik} *and* ε_{ij} *are the elasticities of output with respect to each potential* $K + N$ *inputs, and* s_{ik} *and* s_{ij} *are the respective revenue shares for each input.*

$dlnF_i$ and $dln\omega_i$ denote the growth rates in fixed costs and technical efficiency, with the base given by Q_i, both of which are weighted by total plant-level revenue divided by aggregate value-added, known as the Domar weight. The gaps are now between the output elasticities and the revenue shares. They are multiplied by the Domar weight in aggregation.

When intermediate inputs are separable in the gross output production function, we can represent the growth rate formulation in terms of the value-added production function (see e.g. Bruno (1978)). In this case, we can express the decomposition as

$$\sum_i D_i^v dln\omega_i^v + \sum_i \sum_k D_i^v (\varepsilon_{ik}^v - s_{ik}^v) dlnX_{ik} + \sum_i \sum_j D_i^v (\varepsilon_{ij}^v - s_{ij}^v) dlnM_j^i - \sum_i D_i^v dlnF_i^v, \qquad (10)$$

where the Domar weight is equal to the plant's share of value-added $D_i^v = \frac{VA_i}{\sum_i VA_i}$, the shares become value-added revenue shares (e.g. $s_{ik}^v = \frac{W_{ik} X_{ik}}{VA_i}$). The elasticities are now those for the value-added production

[6] A large literature on adjustment costs derives implications of the adjustment costs for the input demand equations directly (see Bentolila and Bertola (1990) and the review in Bond and Reenan (2008)).

function, which are equal to

$$\varepsilon_{ij}^{v}=\frac{\varepsilon_{ij}}{1-s_{im}},$$

the elasticities from the gross output production function divided by $(1-s_{im})$, with $s_{im}=\frac{\sum_j P_j M_{ij}}{P_i Q_i}$, the ratio of intermediate expenditures to revenues. Similarly, $lnF_i^v=\frac{lnF}{1-s_{im}}$. The value-added technical efficiency shock is derived from the value-added production function, and the Cobb-Douglass case has

$$ln\omega_i^v=ln(VA_i)-\beta_0^v-\sum_k \varepsilon_{ik}^v lnX_{ik}$$

with β_0^v the intercept of the value-added production function.[7] The relationship between the value-added technical efficiency shock and the gross output production function technical efficiency shock is

$$ln\omega_i^v=\frac{ln\omega_i}{1-s_{im}}.$$

In (9) and (10) the Domar weight is applied to both the growth rates in plant-level technical efficiency and reallocation wedge terms. The relationship between these two weights for the gross output and the value-added production function specifications is given by Corollary 2.

Corollary 2

$$\sum_{i=1}^{N} D_i^v=1 \ \ and \ \ \sum_{i=1}^{N} D_i\geq 1,$$

with

$$\sum_{i=1}^{N} D_i=1$$

if and only if

$$dM_{ij}=0 \ \ \forall i,j.$$

The result follows immediately from the definitions of the weights. Corollary 2 shows that in any economy with intermediate inputs, the Domar weights for the gross output production function will sum to a quantity that is strictly greater than one (Hulten (1978)). Consider some additional output that arises either from reallocation of inputs or plant-level technical efficiency gains. When part of it is used as an intermediate input elsewhere, there is a "ripple-effect". An increase to j in intermediate deliveries leads to more output of j, some of which may go directly to final demand and some of which may be sent off as intermediates to plant i or plant k. When the accounting traces out the final impact of the plant-level gain, the appropriate weight is the Domar weight.

[7] If the elasticity of output with respect to each intermediate input is not equal to the intermediate's revenue share, then the estimated residual will include additional terms that reflect these differences. See Basu and Fernald (1995).

In the case of the value-added production function, the growth rate in the technical efficiency residual is already defined in terms of the plant-level value-added. The value-added Domar weight is the share of the plant's value-added in the aggregate. These weights are shares, and so they sum to one.

The next result provides two conditions under which RE contributes zero to changes in aggregate final demand.

Corollary 3

Two sufficient conditions for RE=0 are
1) there is no change in the allocation of inputs:

$$dX_{ik} = 0 \ \ \forall i, k,$$

and

$$dM_{ij} = 0 \ \ \forall i, j,$$

or
2) the "Neoclassical"conditions hold:

$$P_i \frac{\partial Q^i}{\partial X_k} = W_{ik} \ \ \forall i, k,$$

and

$$P_i \frac{\partial Q^i}{\partial M_j} = P_j \ \ \forall i, j.$$

Corollary 3 follows directly from the expression for RE and shows that there are two polar opposite settings in which reallocation as we have defined it can be zero. Case 1 above would occur in the extreme case when frictions or adjustment costs are so high that no primary inputs move between plants and no intermediate inputs are exchanged. Aggregate final demand will increase if inputs started moving from the low value to higher value activities and RE would reflect these movements as APG increased.

Case 2 is the "Neoclassical"setting, where marginal revenue products equal input prices across all inputs and plants. Inputs are being continuously reallocated across plants at all times in response to changes in economic conditions to maintain these conditions. If frictions were introduced that prevented the economy from achieving this "Neoclassical"condition, then the RE measure would reflect the lost output as the economy departed from its production possibilities fronties.

Assuming fixed and sunk costs are always positive, we can bound from below the amount of growth arising from reallocation.

Corollary 4

If $F_i \geq 0 \ \ \forall i$, then

$$RE \geq APG - TE.$$

The result follows from (8). Aggregate reallocation defined by RE must be greater than APG minus the change from technical efficiency. This bound can be useful for a researcher wanting a quick indicator of whether reallocation significantly impacts APG.

4. Discussion

In this section, we discuss how to construct counter-factuals, how types of savings that include investment, net exports, and inventories impact APG, and finally APG discontinuities.

Counterfactuals

We consider the movement of the economy from an environment $E0$ to $E1$. We assume we can index the path of the movements of inputs, outputs and prices between $E0$ and $E1$ over the interval $t \in [0, 1]$. Given the specifics of the transition, the change in aggregate productivity growth is just the integral over the path

$$\Delta APG \equiv \int_0^1 \sum_i P_{it} dY_{it} - \int_0^1 \sum_i \sum_k W_{kt} dX_{ikt}.$$

Suppose one wanted to measure allocative efficiency gains across two economic environments $E0$ and $E1$, holding constant changes in technical efficiency and fixed and sunk costs. Measurement of the costs of increasing or decreasing gaps in the economy has been the subject of much research, including Hopenhayn and Rogerson (1993), Petrin and Sivadasan (2010), Hsieh and Klenow (2009), and Restuccia and Rogerson (2007).[8] For example, $E0$ might denote the state of the economy with firing costs and $E1$ might denote the economy after all firing costs have been eliminated. We use the reallocation terms to define the change in aggregate productivity growth due to changes in allocative efficiency:

$$\Delta AE \equiv \int_0^1 \sum_i \sum_k (P_{it} \frac{\partial Q_{it}}{\partial X_k} - W_{kt}) dX_{kt} + \int_0^1 \sum_i \sum_j (P_{it} \frac{\partial Q_{it}}{\partial M_j} - P_{jt}) dM_j$$

The contribution of allocative efficiency is equal to the extra value of output going to final demand net of any extra primary input costs that occurs as inputs are reallocated in the move from $E0$ to $E1$.[9]

As a simple example, consider the case of a single (labor) input firm. Suppose the firm starts from an economic environment ($E0$) where the firm has a gap positive between the VMP for labor and the wage. Assume the gap is eliminated in the new environment $E1$ by adding labor. The allocative efficiency gain is just the area under the VMP curve and above the wage curve between the level of labor in $E0$ and $E1$.

[8] For example, in the context of plant-level data Petrin and Sivadasan (2010) use RE to motivate their calculation of the cost of increasing severance pay in Chile, which increased from zero in 1984 to one month wages per year of employment (up to a maximum of 11 months wages).

[9] Another variant would add back the adjustment costs associated with these input movements, which would be captured in part in $\sum_i P_i dF_i$.

Savings

The Domar weight implicitly assumes that all output goes immediately to consumption, and most growth accounting measures count all of an economy's current production towards growth. However, there is an old question in the growth literature (and the growth accounting literature) regarding how to value output that is not immediately consumed, like investment, net exports, or inventories. A complete treatment of these types of savings requires a model for future consumption and a definition of growth based on expected discounted consumption (as opposed to current consumption). Authors like Weitzman (1976) and Basu and Fernald (2002) provide a set of assumptions under which the value of current savings exactly equals the value of expected discounted consumption. Alternatively, to the extent that plant-level investment, imports and exports, or inventories are observed directly in the data, they can be separated out to see how much growth measures based entirely on current consumption are affected by these different forms of savings.

If all of an economy's current production is counted towards growth, then all investment is included as are all export revenues because they are domestic production. All expenditures on imports are excluded because they are foreign supply. In plant-level data, revenues usually include export sales and similarly plant-level input expenditures include imported inputs. If plants separately record exported output from domestic output and imported inputs from domestic inputs then one can check how big a role they play in calculated APG.

Inventories

We start with output inventories, rewriting output as

$$Q^i = Y_i + Q^i_{inv} + \sum_i M_{ij} + F_i$$

with Q^i_{inv} denoting the amount of current output set aside to inventory. Many growth accounting measures If we count all output at the time of production towards growth in final demand then the change in plant i's contribution to final demand is given by $P_i(dY_i + dQ^i_{inv})$, with output dQ^i_{inv} counted at price P_i.

Q^i_{inv} has neither contributed to intermediate input use nor to final demand when it is added to final demand, and the price at which Q^i_{inv} is valued is a function of output levels excluding these inventories. One may prefer to define APG by counting Q^i_{inv} at the time that it is sold and with those prevailing prices. If inventories are measured, we can adjust estimated APG to follow this convention.

With regard to intermediate input inventories, if sellers record all of their revenues and buyers record all input expenditures, then intermediate inputs that are purchased and then not used in production cancel out in the aggregation across plants. The expenditures on creating the first inputs have been already accounted, but the value of final demand to which they contribute has not entered APG. In this sense it is a type of savings.

Discontinuities

Some kinds of economic phenomena can lead to discontinuities in APG. These jumps can arise when new product development or marketing leads to large discrete increases in sunk or fixed costs before or at the time of introduction. Similarly, adjustment frictions for inputs that lead to zones of inaction typically result in discrete changes in the input level when the plant does adjust (see e.g. Caplin and Krishna (1986) and Bentolila and Bertola (1990)). The costs of the adjustment often take the form of discrete payments at the time of hiring or firing. These discrete changes in inputs can also lead to discrete changes in consumption after agents reoptimize (see Benhabib and Rustichini (1991)). Two issues for APG growth arise in the presence of jumps.

The first issue relates to integration of the APG measure over time, as denoted in (2). When an integrand has discontinuities, one must establish that the integral is well-defined. There is a large literature on jump processes and the important result for the integral of APG aggregate growth is that there are not "too many"jumps in the integrand. Specifically, the integrand must be a *cadlag* function, which is either right- or left-continuous.[10]

A second issue relates to the value of this integral. At the time of a jump, APG may not capture aspects of the economic environment that the researcher may want included in the growth measure. For example, APG will not generally measure the surplus gains or losses that occur exactly at the time of the jump. In the case of a new good, APG will not count the entire surplus gain when it enters at a price below the reservation (zero-demand) price.

In principle it is easy to adjust the APG measure to account for these changes by defining a new variable $S_i(t)$ for every good i which reflects the current surplus associated with the good. Then dS_i will reflect the change in surplus, and this can be added to APG to produce a new aggregate measure. Measuring dS_i is likely to require more than just the available plant-level data (see Petrin (2002) and Goolsbee and Petrin (2004)).

5. Definitions of Aggregate Productivity Growth Based Exclusively on Technical Efficiency

Many definitions of aggregate productivity growth are based solely on weighted sums of plant-level technical efficiency, and thus ignore all of the reallocation terms in the definition of APG from Section 2.[11] Empirical examples of definitions of average productivity growth based only on technical efficiency include Bailey, Hulten, and Campbell (1992), Olley and Pakes (1996), Griliches and Regev (1995a), and Foster et al. (2001), all of which are defined as the change in the first moment of the plant-level distribution of technical

[10] In the literature on "Stochastic Integration with Jumps" processes are typically written as the sum of three terms. The first term is deterministic. The second term is stochastic and continuous, but not differentiable, often written as Brownian motion. The final term is the "jump" term, and is often modeled as Poisson.

[11] Indeed, popular models such as Melitz (2003) owe much of their elegance to the fact that they can derive a rich set of predictions based on a single-factor measure of plant productivity (and the distribution thereof.)

efficiency. In this section, we compare APG with these indices, which we refer to as Bailey, Hulten, and Campbell (BHC) indices.[12]

Define s_i to be the weight in the aggregation to the first moment of the technical efficiency distribution. In practice this weight is almost always either the gross-output share or the labor share for plant i. In continuous time we then write the original BHC productivity growth index as:

$$BHC \equiv d\sum_i (s_i\, ln\omega_i) = \sum_i s_i\, dln\omega_i + \sum_i ln\omega_i\, ds_i. \tag{11}$$

The BHC measure decomposes into the two right-hand-side terms. The first term is referred to as the aggregate technical efficiency term and the second term is called the aggregate reallocation term.

Corollary 5 shows the relationship in growth rates between aggregate APG technical efficiency (APG_{TE}) and aggregate BHC technical efficiency (BHC_{TE}), and provides a way to calculate APG_{TE} from BHC_{TE}.

Corollary 5

Let $s_i = \frac{P_i Q_i}{\sum_i P_i Q_i}$. *Then*

$$APG_{TE} = \frac{1}{\sum_i P_i Y_i} \sum_i P_i Q_i dln\omega_i = \frac{\sum_i P_i Q_i}{\sum_i P_i Y_i} * \sum_i s_i dln\omega_i = \frac{\sum_i P_i Q_i}{\sum_i P_i Y_i} * BHC_{TE}.$$

The only case in which APG and BHC technical efficiency will be equal is when the gross output share is the aggregation weight and there are no intermediate input deliveries in the economy. Otherwise the difference between the two is increasing in the amount of gross output in the economy that is intermediate input. For example, if the value of intermediate inputs accounts for half of the value of gross output, then BHC technical efficiency understates by one-half the contribution of technical efficiency growth to aggregate final demand.

The contribution of reallocation to the growth rate in the BHC framework is given by the last term in (11) while in our framework, it is given by the second and third terms in (9). BHC reallocation is simply the weighted change in shares when the weight is the log of the plant-level measure of technical efficiency. Putting aside differences in weights (Domar weights in APG and revenue or labor shares in BHC), the basic conceptual difference is that BHC reallocation tracks movements across plants with different levels of technical efficiency while APG reallocation tracks movements of inputs across plants with different wedges between the value of the marginal product and marginal cost. In general, wedges and technical efficiency at the plant-level are unrelated to one another as profit-maximizing plants at all levels of technical efficiency try to keep wedges between marginal revenue and marginal cost as small as possible. Put another way,

[12] According to Hulten, the original genesis for the BHC aggregate was not to add up to growth accounting APG. Instead, it was an exploratory test statistic they were using to evaluate whether the U.S. Manufacturing Census was more consistent with a "macro-divisia index model" or the "looser Schumpeterian paradigm of creative destruction," the latter of which was consistent with persistent plant level shocks and reallocation of output across plants over time.

while more technically efficient plants can produce more output given any input level, in equilibrium more technically efficient plants will use more inputs so that marginal revenue is closer to marginal cost.[13]

One consequence of this difference is that the BHC reallocation will generally be non-zero in the canonical "Neoclassical"setting from Corollary 2 where all wedges equal zero.[14] Thus BHC reallocation generally reports changes in aggregate productivity growth when no more growth in final demand can be achieved via the reallocation of inputs. A second consequence is that BHC aggregate productivity growth can be *negatively* correlated with changes in aggregate final demand holding input use constant. We next illustrate this point with a very simple example, and provide a more detailed example based on a Hotelling model in the Appendix.

In the following single-good two-firm economy example APG decreases but BHC increases. Assume firm one is more technically efficient, so for production technologies $Q_1(\cdot)$ and $Q_2(\cdot)$, $Q_1(l) > Q_2(l)\ \forall l$, and suppose both technologies exhibit decreasing returns to scale, defined as $\frac{\partial Q^2_{(\cdot)}(l)}{\partial l^2} < 0\ \forall l$. Labor is supplied inelastically at quantity $\overline{L}$ and the consumption good price is normalized to 1. The output maximizing allocation of labor is given at

$$\frac{\partial Q_1(l_1^*)}{\partial l} = \frac{\partial Q_2(l_2^*)}{\partial l},$$

with $l_1^* + l_2^* = \overline{L}$. For every unit of input reallocated from firm 2 to firm 1 aggregate final demand falls because for any l_1 and l_2 such that $l_1 > l_1^*$ and any $l_2 < l_2^*$

$$\frac{\partial Q_1(l_1)}{\partial l} < \frac{\partial Q_1(l_1^*)}{\partial l} = \frac{\partial Q_2(l_2^*)}{\partial l} < \frac{\partial Q_(l_2)}{\partial l}.$$

By construction APG would *fall* by exactly this amount. In contrast, the indices based entirely on technical efficiency *increase* because inputs and output are moving to the more technically efficient firm.

6. Discrete Time Approximations

Plant-level data are of course aggregated to discrete intervals. In this section, we provide approximations for integrals of our measure of aggregate productivity growth and the elements of its decomposition.

We use Tornquist-Divisia approximations for all of our calculations, and we chain-weight to update prices on an annual basis (they are included in the Domar weights). For example, for the growth rate version of (4), a standard approximation is given as

$$APG_G = \sum_i \overline{D}_{it}^v \Delta lnVA_{it} - \sum_i \overline{D^v}_{it} \sum_k \overline{s}_{ikt}^v \Delta lnX_{ikt} \tag{12}$$

where $\overline{D}_{it}^v$ is the average of plant i's value-added share weights from period t-1 to period t, Δ is the first difference operator from $t-1$ to t, $\overline{s}_{ikt}$ is the average across the two periods of plant i's expenditures for

---

[13] We suspect this difference is the main reason that BHC and APG reallocation diverge so significantly in our data.

[14] See also the critique in Fox (2003).

the kth primary input as a share of plant-level value-added. Note that we do not need production function estimates to calculate (12).

We do need estimates of production function parameters and residuals to estimate the components of the decomposition. Equation (8) can be estimated in discrete-time by:

$$APG_G = \sum_i \overline{D}_{it}\Delta ln\omega_{it} + \sum_i \overline{D}_{it}\sum_k(\varepsilon_{ik} - \overline{s}_{ikt})\Delta lnX_{ikt} + \sum_i \overline{D}_{it}\sum_j(\varepsilon_{ij} - \overline{s}_{ijt})\Delta lnM_{ijt} - \sum_i \overline{D}_{it}\Delta lnF_{it}, \tag{13}$$

where again bars over variables denote two-period averages and Δ is the first-difference operator. We estimate the production function parameters in logs to obtain estimates for ε_{ik} and ε_{ij}. For the growth rate in plant-level technical efficiency, we use the posited functional form for the production function to calculate the residuals, and then take the first difference. For example, if we assume a Cobb-Douglas production function, we would take first differences of an estimate of:

$$ln\omega_i = lnQ_i - (\sum_k \varepsilon_{ik}lnX_{ik} + \sum_j \varepsilon_{ij}lnM_{ij}). \tag{14}$$

If intermediate inputs are separable in the production function then one can approximate the decomposition using a value-added production function to construct estimates of the elasticities and changes in technical efficiency. In this case the decomposition becomes:

$$APG_G = \sum_i \overline{D^v}_{it}\Delta ln\omega^v_{it} + \sum_i\sum_k \overline{D^v}_{it}(\varepsilon^v_{ik} - \overline{s}^v_{ikt})\Delta lnX_{ikt} + \sum_i\sum_j \overline{D^v}_{it}(\varepsilon^v_{ij} - \overline{s}^v_{ijt})\Delta lnM_{ijt} - \sum_i \overline{D^v}_{it}\Delta lnF^v_{it}, \tag{15}$$

with the value-added residual calculated as

$$ln\omega^v_i + \sum_j(\varepsilon_{ij} - s^v_{ijt})lnM_{ijt} = lnVA_i - (\sum_k \varepsilon^v_{ik}lnX_{ik}). \tag{16}$$

Note that we have explicitly included in the residual the terms related to the intermediate inputs, which will be non-zero if the elasticity of output is not equal to the ratio of expenditure on the input to total revenues for one or more intermediate inputs.[15]

Our approach is to compare the APG and BHC indexes using the value-added representation of productivity growth, so we use the value-added residual and apply the value-added share weight, as in (10). Given any estimator of production function coefficients our estimate of plant-level technical efficiency is given as

$$ln\widehat{\omega^v}_{it} = ln(VA_{it}) - \left(\widehat{\beta^v}_j + \widehat{\epsilon^v}_{jP}\, lnL^P_{it} + \widehat{\epsilon^v}_{jNP}\, lnL^{NP}_{it} + \widehat{\epsilon^v}_{jK}\, lnK_{it}\right).$$

where $\widehat{\beta^v}_j$ and $\widehat{\epsilon^v}_{j\cdot}$ denote the estimated intercept and elasticities of value-added with respect to the inputs in industry j. We apply the same weight in the BHC index calculation, thus abstracting from the usual difference between APG and BHC that is induced by BHC using weights that do not aggregate to APG.

[15] The additional term in the residual arises because value-added is defined by subtracting the expenditures on intermediate inputs. When the revenue share equals the elasticity, the intermediate terms cancel out in the move from the gross output production function to the value-added production function.

The BHC measurement with discrete data is

$$BHC_T = \sum_i s_{G_{iT}} ln\omega_{iT} - \sum_i s_{G_{i,T-1}} ln\omega_{i,T-1}. \tag{17}$$

Given the common weights, we can write

$$BHC_T = \sum_i \frac{(s_{it} + s_{i,t-1})}{2} ln(\frac{\omega_{it}}{\omega_{i,t-1}}) + \sum_i \frac{(ln\omega_{it} + ln\omega_{i,t-1})}{2} * \Delta s_i, \tag{18}$$

where a Tornquist approximation is used. The first term is the technical efficiency term and the common weight ensures that BHC and APG agree on aggregate technical efficiency growth. The second term is the BHC reallocation term.

7. Data and Estimation

We use an annual manufacturing census from Chile to illustrate the empirical relevance of the issues we raise. The data span the period 1979 through 1995. They have been used in numerous other productivity studies, and we refer the interested reader to those papers for a more detailed data description.[16]

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 nominal gross output, two types of labor, capital, and intermediate inputs, including fuels and electricity. Because of the way our plant-level data are reported, we treat plants as firms, although there are probably multi-plant firms. Labor is the number of person-years hired for production, and plants distinguish between their blue- and white- collar workers. The method for constructing the real value of capital is documented in Liu (1991).[17]

We use double deflated value-added as our dependent variable, defined as

$$VA_{it} = \frac{P_{it}Q_{it}}{P_{st}} - \frac{\sum_j P_{jt}M_{ijt}}{P_{st}^M},$$

where P_{st} denotes industry s's gross output deflator, with $s(i)$ being the 3-digit industry s of which i is a part, and P_{st}^M is a 3-digit industry price index for materials. Our value-added specifications include three primary inputs as regressors: production workers L_{it}^P, non-production workers L_{it}^{NP}, and capital K_{it}. We posit a Cobb-Douglass production function and we estimate production functions separately for each 3 digit industry code. We compare results across four different production function estimators, including Ordinary Least Squares, fixed effects, Levinsohn and Petrin (2003), and the Wooldridge (2009) modification of the Levinsohn-Petrin estimator. Our preferred estimator is the Wooldridge-Levinsohn-Petrin estimator

[16] For example see Liu and Tybout (1996), Levinsohn (1999), and Levinsohn and Petrin (2003).

[17] 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.

that uses plant-level materials use as the proxy. It corrects for the simultaneous determination of inputs and technical efficiency, does not maintain constant returns to scale or require cost minimization without input adjustment costs to identify production function parameters, is robust to the Ackerberg, Caves, and Frazer (2008) criticism, and is programmed with a single line of code in Stata (available from the authors on request.)[18] Our estimate of plant-level technical efficiency is then:

$$ln\widehat{\omega^v}_{it} = ln(VA_{it}) - \left(\widehat{\beta^v}_j + \widehat{\epsilon^v}_{jP}\, lnL_{it}^P + \widehat{\epsilon^v}_{jNP}\, lnL_{it}^{NP} + \widehat{\epsilon^v}_{jK}\, lnK_{it}\right),$$

where $\widehat{\beta^v}_j$ and $\widehat{\epsilon^v}_{j\cdot}$ denote the estimated intercept and elasticities of value-added with respect to the inputs in industry j.

8. Results

The purpose of this section is to empirically compare different measures of productivity growth using the Chilean data.[19] Our goal is less on understanding what drove productivity growth and more on highlighting the importance of just how that growth is measured.

All of our results are based on plant-level data, and we readily acknowledge that our results may suffer from some of the measurement problems that typically affect studies using plant-level data. For example, we do not observe plant-level output price deflators and use industry deflators instead. Measurement of capital levels and utilization rates are probably problematic as is our insistence on only two types of labor. Of course, measurement issues afflict both the APG and the BHC measures. The many papers that contribute to our understanding of measurement issues in productivity estimation are relevant to our PL measure.[20]

Table 1 compares the growth rate in value-added with APG for all of Chilean manufacturing. As seen in (4), the difference between the value-added and APG is comprised of the changes in expenditures on primary inputs. Hence, column 1 (which gives the change in value-added) minus column 2 (the change in expenditures on unskilled labor) minus column 3 (the change in expenditures on skilled labor) gives APG, up to an adjustment for capital expenditures (on which we do not have good data). The main message of Table 1 is that the APG tracks the changes in value-added reasonably well. This is because, in this data set anyway, changes in expenditures on skilled and unskilled labor are small relative to changes in value-added. The average annual growth in value-added was 4.04 percent while the average annual increase in APG was 3.91 percent.[21] Although this table and those that follow cover all of manufacturing, as noted above, these measures could be computed for any subset of manufacturing plants.

[18] Ackerberg et al. (2008) do not like the overidentification conditions that Levinsohn-Petrin use to test their model specification. Wooldridge (2009) shows when the over-identification moments are dropped from the estimation routine the estimator reduces to a simple IV setup that is one line of code in Stata and is robust to their critique.

[19] We have also computed results using Colombian data. The results are broadly similar and are included in Petrin, Polanec, and Nishida (2010). For brevity's sake, we only report the Chilean results in this paper.

[20] See e.g. Berndt and Fuss (1986), Morrison and Diewert (1990), Hulten (1992), and Morrison (1986).

[21] The striking figures for 1982 coincide with the debt crisis, the decline in copper prices, and the accompanying severe recession that Chile experienced that year.

Table 2 decomposes APG into the component due to technical efficiency and that due to reallocation. This table quantifies equation (10). The first column of Table 2 gives APG. The next three columns report that portion of the overall productivity growth that is due to technical efficiency when the production function is estimated by OLS, fixed effects, and Wooldridge LP respectively. These columns report the first term of (10) for each of the different production function estimators. The last three columns of Table 2 give the portion of productivity growth due to reallocation. This measure of reallocation includes entry and exit (as these are just an extreme form of reallocation.)

Focusing on the average annual changes at the bottom of the table, about half of productivity growth is due to technical efficiency depending on the production function estimator. Reallocation accounts for the rest of productivity growth. A key point here is that APG reallocation is almost always positive. (The 1982 recession when resources left the economy and the price index plummeted is the obvious exception.) This generally positive contribution of reallocation to productivity growth makes perfect economic sense. Resource reallocation *ought* to contribute positively in any reasonably well-functioning economy, or inputs are systematically reallocating from higher-value to lower-value activities.

Tables 3 and 4 compare the widely used BHC index of productivity growth to the APG index. Table 3 first compares APG (reported in column 1 to make comparisons easy) to four estimates of of BHC productivity growth. The BHC productivity index depends on the production function estimator and the table reports the BHC index for four estimators. The contrast between BHC and APG is pretty stunning. While on average (the bottom rows) the APG and BHC estimates give similar average annual productivity growth rates (3.9 versus 5.2 to 6.3 respectively), the variance of the BHC productivity growth rates is enormous. This is true regardless of how the production function is estimated.

Table 4 shows why the APG and BHC indices so differ. We focus in Table 4 on results using the Wooldridge estimator for the production function. Columns 1 and 2 give the give the previously reported APG and BHC productivity indices respectively. We then report the decomposition of each index into technical efficiency and reallocation. In table 4, we use value-added shares as weights for both the APG and BHC indices in order to highlight what is driving the difference between them. With value-added shares as weights, APG and BHC measure the contribution of technical efficiency identically. The third column reports this contribution. (It is the same as column 4 of Table 2.) The last two columns report the BHC and APG estimates of reallocation respectively. There are two key points. First, BHC reallocation is negative about half the time (and usually largely so.) Second, the BHC estimate of reallocation is typically huge compared to the APG estimate in a given year. At the extreme, the BHC estimate is about 320 times larger (and of opposite sign) than the APG estimate (in 1980) but ratios of 1-20 times larger occur in about half the years. The very large role for reallocation found with the BHC index has spurred a great deal of innovative and influential theoretical modelling (Melitz (2003)) and ensuing empirical studies. That motivating role may well be an artifact of the BHC index. Put another way, it's unclear that the *relatively* modest role for reallocation with the APG index and given in column 5 would have generated the burgeoning literature that the BHC index prompted.

9. Conclusions

We have shown how to extend the traditional definition of aggregate productivity growth to plant-level data. Specifically, we show how to aggregate plant-level data to APG, and how to use plant-level data to decompose APG into several terms, including a term that aggregates changes in plant-level technical efficiency, and several more terms related to the reallocation of inputs across plants. The extension requires us to confront several "non-neoclassical" features that impact plant-level data such as plant-level heterogeneity, the entry and exit of goods, fixed and sunk costs, market power, and adjustment costs like hiring, firing and search costs.

Our measure of aggregate productivity growh has several attractive features. It is in the spirit of the history of productivity growth measures that use aggregate data to track final demand and input costs. It is also robust to kinks and jumps in APG, which are frequently generated by theoretical models of growth. Our measure also provides practitioners with a measurement useful for cost-benefit/policy analysis that is readily comparable across time, industries, countries, and empirical studies.

Our measures of the roles of technical efficiency and reallocative efficiency (in (5) and (6)) provide the basis for the calculation of the impact of the frictions on aggregate final demand. Measurement of the costs of introducing wedges into the economy has been the subject of much research, including Hopenhayn and Rogerson (1993), Petrin and Sivadasan (2010), Hsieh and Klenow (2009), and Restuccia and Rogerson (2007). In order to determine the cost to final demand from the allocative inefficiency(ies), our setup shows that the researcher must determine how inputs would have been distributed but for the policy, as in Hsieh and Klenow (2009) and Petrin and Sivadasan (2010). Given the "but for" distribution of inputs, (8) shows how to calculate the cost to final demand arising from the distortions.

We compare APG to several variants of the widely-used indexes proposed in Bailey et al. (1992), including those in Olley and Pakes (1996), and Foster et al. (2001). While these indexes have decompositions into technical efficiency and reallocation terms that are similar in spirit to those we develop, economic theory indicates that the BHC measure will depart from APG on both the technical efficiency and reallocation dimensions. Perhaps the biggest difference is on the definition of reallocation. APG weights input movements using differences in the gaps between marginal revenue products and input prices, as in (8). BHC weights input movements using differences in technical efficiency across plants, as in (18). In equilibrium, plants choose input levels to equate expected marginal revenue with expected cost of the input, regardless of their level of technical efficiency. Thus the BHC measure uses no information on the differences between marginal revenue products and input prices in its assessment of growth arising from reallocation. Our theoretical findings are consistent with our empirical findings for manufacturing industries in Chile, where the BHC index and its reallocation component behave erratically relative to the growth in technical efficiency and value-added.

References

Ackerberg, D., Caves, K., and Frazer, G. (2008). Structural identification of production functions. Working Paper.

Aghion, P., and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323–351.

Aghion, P., and Howitt, P. (1994). Growth and unemployment. *Review of Economic Studies*, *61*(3), 477–494.

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.

Basu, S., and Fernald, J. (1995). Are apparent productive spillovers a figment of specification error. *Journal of Monetary Economics*, *36*(1), 165–188.

Basu, S., and Fernald, J. (2002). Aggregate productivity and aggregate technology. *European Economic Review*, *46*, 963–991.

Benhabib, J., and Rustichini, A. (1991). Vintage capital, investment, and growth. *Journal of Economic Theory*, 323–339.

Bentolila, S., and Bertola, G. (1990). Firing costs and labour demand: How bad is eurosclerosis?. *Review of Economic Studies*, *57*(3), 381–204.

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.

Bond, S., and Reenan, J. V. (2008). *Handbook of Econometrics: Micro-econometric Models of Investment and Employment*. North-Holland.

Boskin Commission (1996). Toward a more accurate measure of the cost of living. Tech. rep., Social Security Administration Reports and Studies.

Bresnahan, T., and Reiss, P. (1991). Entry and competition in concentrated markets. *Journal of Political Economy*, *99*(5), 997–1009.

Bruno, M. (1978). *An Analysis of the Concept of Real Value-Added, in Production Economics: A Dual Approach to Theory and Applications*, chap. III.1. North-Holland.

Caballero, R., and Engel, E. (1993). Microeconomic adjustment hazards and aggregate dynamics. *Quarterly Journal of Economics*, *108*(2), 359–383.

Caballero, R., and Engel, E. (1999). Explaining investment dynamics in u.s. manufacturing: A generalized (s,s) approach.. *Econometrica*, *67*(4), 783–826.

Caballero, R., and Hammour, M. (1994). The cleansing effect of recessions. *American Economic Review*, *84*(5), 1350–1368.

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.

Caplin, A., and Krishna, K. (1986). A simple dynamic model of employment. Mimeo, Harvard University, Cambridge, Mass.

Domar, E. D. (1961). On the measurement of technological change. *The Economic Journal*, *71*(284), 709–729.

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

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

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. (1995a). Firm productivity in israeli industry 1977-1988. *Journal of Econometrics*, *65*.

Griliches, Z., and Regev, H. (1995b). 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.

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

Hsieh, C.-T., and Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, *74*(4), 1403–1448.

Hulten, C. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, *45*(3), 511–518.

Hulten, C. (1992). Growth accounting when technical change is embodied in capital. *American Economic Review*, *82*(4), 964–980.

Lentz, R., and Mortensen, D. (2008). An empirical model of growth through product innovation. Working Paper.

Levinsohn, and Petrin (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*.

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

Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*.

Morrison, C. (1986). *Capacity Utilization and Productivity Measurement: An Application to the U.S. Automobile Industry*. Kluwer Nijhoff, Boston.

Morrison, C., and Diewert, E. (1990). New techniques in the measurement of multifactor productivity. *The Journal of Productivity Analysis*.

Mortensen, D., and Pissarides, C. (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*.

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

Petrin, A. (2002). Quantifying the benefits of new products; the case of the minivan. *Journal of Political Economy*, *110*, 705–729.

Petrin, A., Polanec, S., and Nishida, M. (2010). Explaining reallocation's apparent negative contribution to growth in deregulation-era developing countries. University of Minnesota Working Paper.

Petrin, A., and Sivadasan, J. (2010). Estimating lost output from allocative inefficiency, with application to chile and firing costs. University of Minnesota Working Paper.

Restuccia, D., and Rogerson, R. (2007). Policy distortions and aggregate productivity with heterogeneous plants. Working Paper 283, Department of Economics, University of Toronto.

Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics*, *39*(3), 312–320.

Trajtenberg, M. (1989). The welfare analysis of product innovations, with an application to computed tomography scanners. *Journal of Political Economy*, *94*, 444–479.

Weitzman, M. (1976). On the welfare significance of national product in a dynamic economy. *Quarterly Journal of Economics*, *90*(1), 156–162.

Wooldridge, J. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economic Letters*, *104*, 112–114.

Appendix

Proof of Lemma 1

Multiply $dY_i = dQ_i - \sum_j dM_{ji}$ by P_i and aggregate over all i to get an expression for the change aggregate value-added:

$$\sum_i P_i dY_i = \sum_i P_i dQ_i - \sum_i P_i \sum_j dM_{ji}. \tag{19}$$

Totally differentiate $Q_i = Q^i(X_i, M_i, \omega_i) - F_i$ and multiply by P_i to get

$$P_i dQ_i = \sum_k P_i \frac{\partial Q_i}{\partial X_k} dX_{ik} + (\sum_j P_i \frac{\partial Q_i}{\partial M_j} dM_{ij}) + P_i d\omega_i - P_i dF_i.$$

Substituting into (19), noting that $\sum_i P_i \sum_j dM_{ji} = \sum_j P_j \sum_i dM_{ij}$, and deducting $\sum_i \sum_k W_{ik} dX_{ik}$, and rearranging yields the decomposition of PL in levels. Translating to growth rates follows immediately. #

A Second Theoretical Example of BHC/PL Divergence

Consider a setting with heterogeneous consumers and imperfect competition. We examine the change in welfare as we shift output from the less technically efficient to the more technically efficient plant. With heterogeneity in the valuation of output, the gains from the cost savings of shifting production to the more efficient plant is eventually offset by the loss in consumer surplus induced by shifting marginal consumers away from the less efficient plant's output, which they value more than the output from the more efficient plant.

We use a classic Hotelling line model, which we can view as a world with two identical goods plus consumer transportation costs or a world with two differentiated products with consumers heterogeneous in tastes. We define total surplus as consumer surplus plus producer surplus. Lemma A1 shows the tradeoff between cost savings and consumer surplus induced by a tax (subsidy) on the less (more) efficient producer.

Lemma A1

Assume consumers are distributed uniformly on the unit interval with x denoting their location on [0,1], with plant 0 (good 0) located at the left endpoint and plant 1 (good 1) at the right endpoint. Let t be the transport cost and τ be the tax on good 0. Denote plant i's technical efficiency c_i, the unit cost of production, and assume these two firms compete Bertrand-Nash in prices. Assume a consumer located at x values good 0 at

$$q^* - p_0 - t * x$$

and good 1 at

$$q^* - p_1 - t * (1 - x).$$

Letting $x^*(\tau)$ denote the indifferent consumer at tax τ, total surplus W at tax τ is equal to

$$W(\tau) = q^* - \frac{1}{2}t - c_1 + (c_1 - c_0)x^*(\tau) + tx^*(\tau)(1 - x^*(\tau)).$$

The derivative of surplus with respect to τ is

$$\frac{\partial W(\tau)}{\partial \tau} = (c_1 - c_0)\frac{\partial x^*(\tau)}{\partial \tau} + t * (1 - 2x^*(\tau))\frac{\partial x^*(\tau)}{\partial \tau}.$$

The tax τ^* that maximizes surplus is

$$\tau^* = 2(1 - \frac{c_1}{c_0}).$$

See proof below.

For any level of τ at which both plants produce the indifferent consumer is given by

$$x^* = \frac{1}{2} - \frac{1+\tau}{6t}c_0 + \frac{1}{6t}c_1.$$

$\frac{\partial x^*}{\partial \tau} < 0$ as an increase in the tax induces an output shift from plant 0 to plant 1. For any level of τ at which both firms are producing, a small increase in the tax leads to a change in total costs to the economy of $(c_1 - c_0)\frac{\partial x^*}{\partial \tau}$. We assume $c_0 > c_1$ which implies the tax must be non-negative to decrease total costs.

The negative impact of the tax is that it distorts consumption. Consumer surplus falls by $t * (1 - 2x^*(\tau))\frac{\partial x^*}{\partial \tau}$, which is the difference in valuation of the marginal consumer $x^*(\tau)$, for whom the additional output from plant 1 is valued less than the lost output from plant 0. The optimal tax τ^* trades these margins off and is chosen to equate

$$c_0 - c_1 = t * (1 - 2x^*(\tau^*)).$$

For $\tau \in [0, \tau^*)$, increases in the tax increase total surplus as the lost consumer surplus is less than the cost savings. Both APG and BHC will increase on this range, although only APG will exactly equal the change in total surplus as BHC does not account for the lost consumer surplus. Once $\tau > \tau^*$, total surplus falls as the tax increases because the marginal cost savings are less than the marginal losses in consumer surplus. APG exactly equals the fall in total surplus as the tax increases beyond τ^*. BHC continues to increase even after $\tau > \tau^*$ because it ignores the impact on consumer surplus. Thus BHC is *negatively* correlated with APG over the range for $\tau > \tau^*$.

Proof of Lemma A2

Given prices the indifferent consumer is

$$x^* = \frac{1}{2t}(p_1 - p_0 + t).$$

Plants 0 and 1 choose prices to maximize $(p_0 - c_0)x^*$ and $(p_1 - c_1)(1 - x^*)$ respectively. With $\tau = 0$ equilibrium prices are

$$p_0 = t + \frac{2}{3}c_0 + \frac{1}{3}c_1,$$

and symmetrically for plant 1. In terms of costs, the indifferent consumer becomes

$$x^* = \frac{1}{2} - \frac{1}{6t}c_0 + \frac{1}{6t}c_1.$$

If we apply a tax τ to good zero then equilibrium prices are given as

$$p_0 = t + \frac{2(1+\tau)}{3}c_0 + \frac{1}{3}c_1,$$

$$p_1 = t + \frac{1+\tau}{3}c_0 + \frac{2}{3}c_1,$$

and

$$x^* = \frac{1}{2} - \frac{1+\tau}{6t}c_0 + \frac{1}{6t}c_1.$$

With tax revenues added back to consumption total surplus is given as

$$W = \int_0^{x^*}(q^* - p_0 - tx)dx + \int_{x^*}^{1}(q^* - p_1 - t(1-x)dx + (p_0 - c_0)x^* + (p_1 - c_1)(1 - x^*),$$

which gives

$$W = q^* - \frac{1}{2}t - c_1 + (c_1 - c_0)x^* + tx^*(1 - x^*).$$

$\frac{\partial W}{\partial \tau}$ and the optimal tax follow immediately. #

Entry and Exit

Entry and exit are naturally accommodated in our approach because we sum over all outputs and inputs associated with any good that may at any time be produced. Entry in our setup could be the development of a new good, the replication by a new plant of an existing good, or the reintroduction of a good that at some point exited. Plants may incur different costs for each of these types of entry. For an entirely new good there may be sunk or fixed costs associated with developing the good, which may be incurred before, during, or after the time of production and introduction. These may include one-time research and development costs or the costs associated with some type of demand or marketing analysis. A good that has been developed already but is not currently being produced may have fixed costs associated with keeping open the option of future production, even if no inputs into direct production are being incurred. If one plant copies another plant, entering with an identical (or similar) good, they may only incur some of the sunk costs that the initial developer incurred. Our approach nests all of these situations.

We describe an example of the life-cycle APG accounting of a product i that enters at time t_1 and exits at time t_2. Prior to t_1 when i is not being produced, $Q_i = 0$ so $dQ_i = 0$ (and thus $dM_{ji} = 0 \; \forall j$). If no fixed or sunk costs are being incurred ($F_i = 0$) then the contribution of i to final demand is zero because $dY_i = dQ_i - \sum_j dM_{ji} - dF_i = 0$. If no primary inputs are used, $dX_{ik} = 0 \; \forall k$, and the contribution from i to APG prior to t_1 is zero.

When entry occurs $dQ_i(t_1) > 0$. This output may go to final demand, intermediate input use, or both, and will be reflected in APG in dY_i and dM_{ji}. If there are fixed or sunk costs associated with (re)starting production, then $dF_i > 0$. Increases in primary and intermediate input use will be reflected in dM_{ij} for i and dX_{ik}. When the good exits $dQ_i(t_2) < 0$ as production goes to zero. After exit $Q_i = 0$ so $dQ_i = 0$. Typically primary and intermediate input use will also go to zero unless the plant is incurring costs to keep the option of future production open or the plant has continued obligations to labor or capital.

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

Comparing the Rate of Growth in Aggregate Value Added with APG, 1980-1995

Year	APG=Change in Value Added - Change in Labor Costs			
	Value Added	Unskilled Labor	Skilled Labor	APG
1980	1.63	-0.21	0.04	1.81
1981	4.70	-1.01	-0.46	6.17
1982	-27.38	-2.32	-1.88	-23.18
1983	2.79	-0.44	-0.51	3.74
1984	0.13	0.61	0.09	-0.57
1985	9.70	0.37	0.25	9.07
1986	11.57	0.09	1.03	10.46
1987	1.79	0.55	0.99	0.24
1988	9.56	0.42	0.56	8.58
1989	9.83	0.62	0.59	8.62
1990	2.87	0.16	0.08	2.62
1991	8.28	0.40	0.31	7.57
1992	12.40	0.56	0.42	11.42
1993	6.95	0.19	0.22	6.54
1994	2.66	0.00	0.40	2.26
1995	7.20	-0.02	0.06	7.16
Average	4.04	-.0012	0.14	3.91
Std. Dev.	9.23	0.76	0.68	8.09

TABLE 2

Comparing the Decomposition of APG Across Production Function Estimators

Chilean Manufacturing, 1980-1995

Year	APG	APG Technical Efficiency OLS	APG Technical Efficiency Fixed Effects	APG Technical Efficiency Wooldridge-LP	APG Reallocation OLS	APG Reallocation Fixed Effects	APG Reallocation Wooldridge-LP
1980	1.81	3.61	3.55	2.06	-1.80	-1.74	-0.25
1981	6.17	8.95	7.15	5.47	-2.78	-0.98	0.70
1982	-23.18	-13.04	-15.07	-14.97	-10.14	-8.11	-8.22
1983	3.74	-0.26	0.65	-3.05	4.01	3.10	6.80
1984	-0.57	-3.61	-4.46	2.04	3.04	3.90	-2.61
1985	9.07	10.77	16.96	3.88	-1.69	-7.88	5.19
1986	10.46	5.76	7.01	10.81	4.69	3.45	-0.36
1987	0.24	-6.41	-7.66	-6.66	6.65	7.90	6.90
1988	8.58	2.35	5.34	6.14	6.23	3.24	2.44
1989	8.62	3.46	1.37	0.47	5.16	7.25	8.15
1990	2.62	3.20	4.37	3.07	-0.58	-1.75	-0.45
1991	7.57	3.39	4.70	0.87	4.19	2.87	6.70
1992	11.42	7.19	7.22	6.77	4.23	4.20	4.64
1993	6.54	9.50	4.06	0.72	-2.96	2.48	5.82
1994	2.26	-0.77	-0.99	-0.14	3.03	3.26	2.40
1995	7.16	2.43	3.49	4.87	4.73	3.68	2.29
Average	3.91	2.28	2.35	1.40	1.63	1.55	2.51
Std. Dev.	8.09	6.17	7.17	5.97	4.53	4.58	4.31

TABLE 3

Comparing the Rate of Growth in Aggregate Value Added, APG and BHC

Chilean Manufacturing, 1980-1995

Year	APG	BHC Productivity Growth Across Estimators			
		OLS	Fixed Effects	Levinsohn-Petrin (LP)	Wooldridge-LP
1980	1.81	39.04	72.80	79.35	82.53
1981	6.17	34.34	-1.23	-29.52	-34.91
1982	-23.18	-25.12	13.14	44.44	28.21
1983	3.74	2.47	25.27	14.23	5.77
1984	-0.57	-10.29	-22.68	-19.23	-9.48
1985	9.07	11.19	31.82	38.57	53.71
1986	10.46	7.95	-1.89	-2.28	-9.43
1987	0.24	-0.26	-8.16	-11.80	-3.25
1988	8.58	-0.24	13.06	14.46	27.20
1989	8.62	8.22	-6.48	-13.16	-3.91
1990	2.62	8.63	-1.69	-5.99	-14.03
1991	7.57	6.77	2.95	-2.67	-30.45
1992	11.42	13.98	-0.07	-8.69	14.81
1993	6.54	10.81	-19.18	-40.24	-47.84
1994	2.26	-5.55	-3.24	-3.99	-4.09
1995	7.16	-19.14	5.98	33.10	37.37
Average	3.91	5.17	6.27	5.41	5.76
Std. Dev.	8.09	16.54	22.68	30.67	33.87

BHC is calculated using production function parameters that vary across 3-digit ISIC estimates using the above estimators. Levinsohn-Petrin is the Levinsohn and Petrin (2003) estimator, and Wooldridge-LP is the Wooldridge estimator which modifies LP by dropping the contemporaneous labor moments so one uses lagged inputs for identification instead of current inputs.

TABLE 4

Comparing the Rate of Growth in Aggregate Value Added with APG, and BHC, including Decomposition, 1980-1995

Wooldridge-Levinsohn-Petrin Estimator

Year	APG	BHC	Technical Efficiency	BHC Reallocation	APG Reallocation
1980	1.81	82.53	2.06	80.47	-0.25
1981	6.17	-34.91	5.47	-40.38	0.70
1982	-23.18	28.21	-14.97	43.18	-8.22
1983	3.74	5.77	-3.05	8.82	6.80
1984	-0.57	-9.48	2.04	-11.52	-2.61
1985	9.07	53.71	3.88	49.84	5.19
1986	10.46	-9.43	10.81	-20.24	-0.36
1987	0.24	-3.25	-6.66	3.41	6.90
1988	8.58	27.20	6.14	21.06	2.44
1989	8.62	-3.91	0.47	-4.38	8.15
1990	2.62	-14.03	3.07	-17.10	-0.45
1991	7.57	-30.45	0.87	-31.32	6.70
1992	11.42	14.81	6.77	8.04	4.64
1993	6.54	-47.84	0.72	-48.56	5.82
1994	2.26	-4.09	-0.14	-3.95	2.40
1995	7.16	37.37	4.87	32.50	2.29
Average	3.91	5.76	1.40	4.37	2.51
Std. Dev.	8.09	33.87	5.97	34.65	4.31

Results include the entire Chilean manufacturing census. BHC is calculated using production function parameters that vary across 3-digit ISIC estimates using the Wooldridge-LP estimator (see paper).