Production and Cost Functions

Allan Collard-Wexler

January 2, 2012

1 Introduction

Production Functions are indispensable tools for Empirical I.O. and Economics in general. Recently there has been considerable progress in estimation techniques which take into account the fact that profit maximizing firms are making input and exit decisions based on their current levels of efficiency.

- **Sources of Aggregate Productivity Growth:** By definition, aggregate productivity growth is due to individual plants within a country becoming more productive. Does productivity increase due to the entry of efficient producers or because productive plants grow more than unproductive plants?

- **Trade:** When a country opens up to international trade this forces plants to compete with firms in other countries. This has the effect of inducing inefficient producers to exit the industry and other firms to raise their productivity. Pavcnik (2002) finds that when trade barriers were lifted in Chili, plant productivity showed a marked increase.

- **Mergers and Monopoly:** When two firms merge are there efficiency gains that will lead to increased productivity? To answer this question we need to estimate returns to scale for the industry to see if larger plants also end up being more productive.

- **Impact of I.T. on Productivity Growth:** There has been sustained discussion on which factors led to the increase in productivity growth during the late 90's. One conjecture is that increase spending on information technology was responsible for this jump in productivity. Looking at the correlation between increases in productivity and plant level spending on I.T. (such as in Athey and Stern (2002)) could allow us to answer this question.

2 Functional Forms and Unobservables

Let's start with the typical Cobb-Douglas production function:

$$Y_{it} = A_{it} K_{it}^{\beta_k} L_{it}^{\beta_l} \tag{1}$$

where Y is value-added, and L and K are salaries and total assets respectively. Taking logs we get the following log-linear form:

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \epsilon_{it} \tag{2}$$

where ϵ is a unobservable input/output. Note that this formulation assume two things. First, the Cobb-Douglas production function assume unit elasticity of substitution between capital and labour. Second, the unobservable ϵ is additively separable in logs and is thus Hicks Neutral, i.e. firms do not differ in their capital to labour ratios (assuming they pay the same rental rate for capital and the same wages for labour), and firms do not vary in their returns to scale.

There are of course other functional forms we could assumed:

- **Translog:**

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_{kk}(k_{it})^2 + \beta_{ll}(l_{it})^2 + \beta_{lk} l_{it} k_{it} \tag{3}$$

- **CES:**

$$Y_{it} = A(\beta_k (K_{it})^{\rho} + \beta_l (L_{it})^{\rho})^{\frac{1}{\rho}} \tag{4}$$

- Another alternative is to assume less functional form such as monotonicity, i.e. $l' < l \Rightarrow y' < y$ and $k' < k \Rightarrow y' < y$ or concavity, linear homogeneity or any assumptions on the functional form of the production function derived from theory.

2.1 Unobservables and Returns to Scale

In Economics there are several different assumptions we can make about unobservables:

1. **Measurement Error:** The unobservables ϵ's are measurement error which is uncorrelated with the x's. This is the usual assumption discussed in Econometrics classes. However, in empirical microeconomics we don't usually perform experiments to collect data. As such, we don't control the factors that may enter into the firm's decision such as the amount of managerial capital that the firm uses or the presence of an unobserved component to productivity. So the assumption that the unobservables are simply measurement error is often hard to swallow.

2. **Expectation Error:** The unobservables can also be due to errors on the part of agents. So for instance, I may observe different firms making decisions on how much capital and labour to purchase in this period, even though I believe that all firms face the same input and product market and use the same technology. Certain firms will purchase more capital and labour than others because they make a mistake when choosing their optimal use of inputs, i.e. $|\frac{\partial \Pi}{\partial L} - w| < \epsilon$ where ϵ represents my inability to verify if I am really at a zero of my first-order condition. This type of error could be quite reasonable in cases where agents face very difficult optimization problems which have a very flat first-order conditions, i.e. if I make a small mistake in evaluating my first order condition, this could lead to a big difference in the amount of capital or labour I choose to use.

3. **Private Information Unobservables:** Sometimes it is useful to assume that each firm knows their own ϵ_i but that other firms don't know about it. It is typically a more reasonable assumption for unobservables, that econometrician (a.k.a. you) knows less about the players in the market than they know about themselves. However, in the case of private information, I also assume that I know as much about about a firm's rivals as the firm itself. There are some cases where this is a reasonable assumption, such as in auctions on Ebay where I know that bidders are anonymous. However, when I model GM's decision to price it's cars, it is unreasonable to assume that Ford would not know more about this unobservable ξ's than I do.

4. **Common Knowledge Unobservables to the Econometrician:** In BLP, firms do not react just to $X\beta - \alpha p$ but to $X\beta - \alpha p + \xi$ which includes the unobserved product characteristics ξ. Thus I assume that all firms know each other's ξ_{-i} when making their decision. Moreover, it is hard to imagine many situation where firms where firms have some uncertainty over their rivals's ξ since they can learn about these after repeated play over many periods.

One example where unobservables play a large role is in estimating returns to scale. For many different policy applications it is important to know if an industry has increasing returns to scale. For example, if two firms merge will prices increase because of greater market power, or decrease because lower marginal costs? However, most production function regressions find *decreasing* returns to scale, which is puzzling given that we also find very large plants in most industries.

Suppose I estimate the following production function:

$$\begin{aligned} q_i^t(\text{cubic yards of concrete}) = \; & \beta_l l_i^t(\text{salaries}) + \beta_k k_i^t(\text{capital}) \\ & + m_i^t(\text{materials}) + A^t + \rho_i^t \end{aligned} \tag{5}$$

Table shows production function regressions using alternatively value added, shipments and cubic yards of concrete as a measure of output. As you can see, in all three regression the return to scale coefficient (the sum of β_l, β_k and β_m) is less than 1. However this makes no sense since Table **??** shows that in larger markets you get bigger plants.

What is most likely going on is the presence of unmeasured inputs into the production process such as managerial ability (henceforth denoted at *man*). If we omit *man* from the production function, and if other inputs such as capital or labor are not perfectly correlated with *man* then this will bias estimates of returns to scale downwards. Formally:

$$\begin{aligned} y_{it} &= \beta_k k_{it} + \beta_l l_{it} + \beta_{man} man_{it} + \epsilon_{it} \\ &= (\beta_k + \beta_{km}) k_{it} + (\beta_l + \beta_{lm}) l_{it} + (\beta_{man} - \beta_{km} - \beta_{lm}) man_{it} + \epsilon_{it} \end{aligned} \tag{6}$$

where $man_{it} = \beta_{km} k_{it} + \beta_{kl} l_{it} + u_{it}$.

2.2 Market Power Versus Productivity

Another issue brough up by Syverson (2004) and de Loecker (2006) is the problem of distinguishing market power from productivity. In most production function regressions, output is measured in dollars instead of in quantities. Thus it is impossible to know if firms are more productive or do they simply exploit their market power to charge higher prices?

As is de Loecker (2006), since revenue are $R_{it} = P_{it}Q_{it}$, then the production function regression estimated using sales has the following bias:

$$r_{it} = \beta_l l_{it} + \beta_k k_{it} - p_{it} + \epsilon_{it} \tag{7}$$

Therefore, firms which can charge higher prices will have upward bias in their measured productivity.

3 Selection and Endogeneity

If I treat the ϵ term as measurement error, then I can obtain consistent estimates with OLS. However these ϵ terms are more appropriately though of as inputs, such as managerial ability, that the firm observes but I don't. Let's rewrite the production function as:

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it} \tag{8}$$

where ω_{it} are productivity factors that the firm observes such as soil quality and η_{it} are productivity factors that the firm does not anticipate such as changes in rainfall from year to year. Anticipated components of productivity are a problem since they will induce a correlation between the firm's choices for capital and labour and the anticipated component of productivity. So for instance the firm's capital choice is determined by:

$$K_{it} = [\frac{p\exp(\omega)\beta_k}{r} L_{it}^{\beta_l}]^{\frac{1}{1-\beta_k}} \tag{9}$$

and thus there is a positive correlation between ω and K induced by the fact that the firm will respond to changes in unobserved productivity. This is called the *endogeneity* bias.

Likewise, the fact that firms choose whetever they will exit in the next period also introduce a *selection* bias. A firm decides on remaining in the industry based on the following decision:

$$\chi(\omega_{it}, k_{it}, x_{it}) = 1\left\{V(\omega_{it}, k_{it}, x_{it}) > \Psi\right\} \tag{10}$$

Now of course, we don't know what ω is, which make this a relationship we can only estimate. If you look at the probits reported in Tables 1 and

2, larger plants and more productive plants are less likely to exit. Thus a big plant can remain in operation with a productivity level that would have caused the smaller plant to exit the industry, and thus $E[\omega k|\chi = 1] < 0$. One way of solving this problem is to use the tradional Heckman Selection approach, with an explicit formulation for the selection equation assuming that ω is normally distributed:

$$\begin{aligned} \chi(\omega_{it}, k_{it}, x_{it}) &= \alpha_k k_{it} + \alpha_x x_{it} + \omega_{it} \\ \Pr[\chi = 1] &= \Phi(\alpha_k k_{it} + \alpha_x x_{it}) \end{aligned} \tag{11}$$

We can estimate this equation as a simple probit. To correct for the selection problem, note that under the assumption that ω has a normal distribution, $E[\omega|\chi_{it} = 1, k_{it}, x_{it}] = \frac{\phi(\hat{\alpha_k} k_{it} + \hat{\alpha_x} x_{it})}{\Phi((\hat{\alpha_k} k_{it} + \hat{\alpha_x} x_{it})}$. So the production function regression becomes:

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \frac{\phi(\hat{\alpha_k} k_{it} + \hat{\alpha_x} x_{it})}{\Phi((\hat{\alpha_k} k_{it} + \hat{\alpha_x} x_{it})} + \eta_{it} \tag{12}$$

3.1 Fixed Effects, Quasi-Differencing

The first application of fixed effect techniques comes from the estimation production functions in agricultural economics. Suppose that the anticipated component of productivity ω is fixed for each firm across time, i.e. $\omega_{it} = \omega_i$. Then if I take first differences, we obtain:

$$y_{it} - y_{it-1} = \beta_k (k_{it} - k_{it-1}) + \beta_l (l_{it} - l_{it-1}) + \eta_{it} - \eta_{it} \tag{13}$$

The only issue is where is the variation in capital and labour usage coming from? In particular, I am tossing out all the cross-sectional variation in capital and labour usage which could help identify coefficients. For example, suppose I am looking at the relationship between studying and exam performance. If I look across individuals I would find that people who study more have higher performance on exams. However, if I look within a class, I would find that exams for which I study less are easier exams on which I will get a higher grade.

If you take a look at Table 3.1, you can see that including fixed effects reduces the capital coefficient dramatically. Thus, eliminating cross-sectional variation reduces the variation in capital that we can use in order to identify β_k.

Alternatively, one can use quasi-differences, so suppose unobserved productivity follows an AR(1) process, i.e. $\omega_{it} = \rho\omega_i^{t-1} + \zeta_{it}$. Then we can take quasi-difference instead of a first-difference:

$$y_{it} - \rho y_{it-1} = \beta_k(k_{it} - \rho k_{it-1}) + \beta_l(l_{it} - \rho l_{it-1}) + \eta_{it} - \eta_{it} + \zeta_{it} \tag{14}$$

Here I can keep more of the cross-sectional variation to identify the parameters of the production function.

3.2 Instrumental Variables

One approach is to find instrumental variable, i.e. a Z such that $E[y|X,Z]$ does not vary with Z, but $E[X|Z]$ does vary with Z. Typically, we use the following moment condition:

$$\begin{aligned} E[\{K,L\}Z] &> 0 \\ E[\epsilon Z] &= 0 \end{aligned} \tag{15}$$

However, in practice it is often difficult to find a good instrument. Certain variables such as the price of goods in other markets may be plausibly uncorrelated with ϵ. However, we have no reason to believe that these Z's will also be correlated with firms capital and labour choices. Syverson (2004) uses local demand shocks for ready-mix concrete as an instrument, since these will be correlated with labour and materials choices (since firms react to changes in demand by lowering their use of inputs) but won't be correlated the underlying unobserved productivity.

3.3 Olley-Pakes

Olley and Pakes (1996) study the effect of opening up the market for long-distance telecommunications by breaking up AT&T on the productivity of telecommunications equipment manufacturers, which went from being a monopoly supplier to operating in a competitive international market.

One of the issues in this litterature is how to deal with the presence of an anticipated unobserved productivity ω. Olley and Pakes (1996) use the fact that investment is increasing in unobserved productivity to proxy for the effect of ω. The timing in this model is as follows: firms need to choose their capital stock one period ahead of production. Then their current period's productivity shock is realized and they choose labour and materials accordingly.

The firm's investment decision is based on the solution to a fairly complicated dynamic programming problem:

$$\max_{\chi \in \{0,1\}} \{\max_i E[V(\delta k + i, \omega', x')|\omega, x], \Psi\} \tag{16}$$

Under certain fairly plausible conditions (covered by Ericson and Pakes (1995)), the investment policy function $i = i(\omega, k, x)$ is stricly increasing in ω. This allows us to invert out the ω term: $\omega_{it} = h(i_{it}, k_{it}, x_{it})$. Now here comes the really interesting part. I don't need to compute the investment policy function from equation 16 and then invert it! Instead I can simply estimate the $h(.,.,.)$ function directly from the data. This insight that policy function can be recovered "non-parametrically" from the data is a very powerfull idea in modern empirical I.O.. This type approach also allows us to recover other objects like bid functions in the context of auctions or investment policy functions in dynamic oligopoly. Oddly enough, it is sometimes easier to estimate a model in modern I.O. than solve their theoretical counterparts since we can recover the firm's policy functions directly.

In the first stage, we estimate the following model:

$$\begin{aligned} y_{it} &= \beta_l l_{it} + \beta_k k_{it} + h(i_{it}, k_{it}, x_{it}) + \eta_{it} \\ &= \beta_l l_{it} + \phi(i_{it}, k_{it}, x_{it}) + \eta_{it} \end{aligned} \tag{17}$$

From this equation you can get estimates of both the labour coefficient $\hat{\beta}_l$ and of the hybrid function $\hat{\phi}(.,.,.)$. In the second stage, note that:

$$\begin{aligned} \omega_{it} &= E[\omega_{it}|\omega_i^{t-1}, k_i^{t-1}, x_{it}] + \zeta_{it} \\ &= g(\omega_i^{t-1}) + \zeta_{it} \\ &= g(h(i_i^{t-1}, k_i^{t-1}, x_i^{t-1})) + \zeta_{it} \\ &= g(\hat{\phi_i^{t-1}} - \beta_k k_i^{t-1}) + \zeta_{it} \end{aligned} \tag{18}$$

where $g(.)$ is simply another one dimensional polynomial function. Suppose that ω was simply an AR(1) process. In this case, the process for ω would simply be:

$$\begin{aligned} \omega_{it} &= \rho\omega_i^{t-1} + \zeta_{it} \\ &= \rho h(i_i^{t-1}, k_i^{t-1}, x_i^{t-1}) + \zeta_{it} \end{aligned} \tag{19}$$

We can substitute equation 25 into the production function regression instead of ω_{it}:

$$y_{it} = \hat{\beta}_l l_{it} + \beta_k k_{it} + g(\hat{\phi_i^{t-1}} - \beta_k k_i^{t-1}) + \zeta_{it} + \eta_{it} \tag{20}$$

This equation can be estimated quite easily.

We can also incorporate selection into the Olley-Pakes framework.

3.4 Levinsohn-Petrin

One of the problems with the Olley and Pakes (1996) setup is the fact that in most years plants make zero investment. To counter this issue, Levinsohn and Petrin (2003) use a different proxy for productivity based on a plant's use of materials. Materials can be used as a proxy since the plant is assumed to order materials once it's productivity shock has been realized, and thus these are strictly endogenous variables.

As before the materials demand equation is:

$$m_{it} = m(k_{it}, \omega_{it}, x_{it}) \tag{21}$$

which can be inverted if the materials decision is strictly increasing in ω_{it}, which is easy to show since the choice of how much materials to purchase is a strictly static decision. To see this, note that the first order condition for materials and labour given a fixed level of capital is:

$$\begin{aligned} \frac{\partial \Pi}{\partial M} &= \beta_m exp(\omega) K_k^{\beta} L_l^{\beta} M^{\beta_m - 1} = p_m \\ \frac{\partial \Pi}{\partial L} &= \beta_l exp(\omega) K_k^{\beta} L^{\beta_l - 1} M_m^{\beta} = w \end{aligned} \tag{22}$$

And thus:

$$\omega_{it} = h(m_{it}, k_{it}, x_{it}) \tag{23}$$

Plug this in to the production function as before:

$$\begin{aligned} y_{it} &= \beta_l l_{it} + \beta_k k_{it} + h(m_{it}, k_{it}, x_{it}) + \eta_{it} \\ &= \beta_l l_{it} + \phi(m_{it}, k_{it}, x_{it}) + \eta_{it} \end{aligned} \tag{24}$$

In the first stage we get estimates for $\hat{\beta}_l$ and $\hat{\phi}_{it}$. As before:

$$\begin{aligned} \omega_{it} &= E[\omega_{it} | \omega_i^{t-1}, k_i^{t-1}, x_{it}] + \zeta_{it} \\ &= g(\omega_i^{t-1}) + \zeta_{it} \\ &= g(h(i_i^{t-1}, k_i^{t-1}, x_i^{t-1})) + \zeta_{it} \\ &= g(\hat{\phi_i^{t-1}} - \beta_k k_i^{t-1}) + \zeta_{it} \end{aligned} \tag{25}$$

Then the second stage estimating equation is:

$$\begin{aligned} y_{it} &= \hat{\beta}_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \omega_{it} + \eta_{it} \\ &= \hat{\beta}_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + g(\hat{\phi_i^{t-1}} - \beta_k k_i^{t-1}) + \zeta_{it} + \eta_{it} \end{aligned} \tag{26}$$

Note that there is a problem with estimating this regression since materials are correlated with the innovation ζ_{it} in today's unobserved productivity ω_{it}. However, we can find an instrument for this regression, which is a Z which is correlated with the dependant variables but uncorrelated with the unobservables: $\zeta_{it} + \eta_{it}$. This is the case for lagged materials since there are uncorrelated with the innovation in productivity, and yet they are correlated with materials and labour through the persistence in ω_{it} and with capital directly since they are correlated directly with ω_i^{t-1}.

3.5 Ackerberg-Caves-Frazer

Ackerberg, Frazer, and Caves (2006) investigate the potential for collinearity in the various proxy for productivity techniques we have been investigating. In particular, once we put in productivity and fixed capital into the regression, why should there be any variation left over from labour and materials?

To see this, note that since the labour decision is purely static, then it depends on the state of the firm $s_{it} = \{\omega_{it}, k_{it}, x_{it}\}$. Thus the labour demand equation is given by:

$$l_{it} = l(\omega_{it}, k_{it}, x_{it}) \tag{27}$$

And likewise for the materials decision:

$$m_{it} = m(\omega_{it}, k_{it}, x_{it}) \tag{28}$$

Thus the production function regression becomes:

$$\begin{aligned} y_{it} &= \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \omega_{it} + \eta_{it} \\ &= \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + h(m_{it}, k_{it}, x_{it}) + \eta_{it} \\ &= \beta_l l(\omega_{it}, k_{it}, x_{it}) + \beta_k k_{it} + \beta_m m(\omega_{it}, k_{it}, x_{it}) + h(m_{it}, k_{it}, x_{it}) + \eta_{it} \end{aligned} \tag{29}$$

Thus it is impossible to separate labour demand, materials demand and unobserved productivity in the first stage. The way this model is speficied, it is not possible for labour or materials to have any variance separate from the firm's state.

Note that we would not have this collinearity problem if labour were also determined a period ahead (say because of a Union's labour contract). However, this would introduce labour into the firm's state space.

Table 3: Production Function Estimates from Pavcnik (2002)*

	OLS	Fixed Effects	Olley-Pakes
Unskilled Labor	0.178 (0.006)	0.210 (0.010)	0.153 (0.007)
Skilled Labor	0.131 (0.006)	0.029 (0.007)	0.098 (0.009)
Materials	0.763 (0.004)	0.646 (0.007)	0.735 (0.008)
Capital	0.052 (0.003)	0.014 (0.006)	0.079 (0.034)

3.6 Productivity Dispersion

To display productivity dispersion more "concretely", I generate the predicted output for each plant as if it were using mean levels of capital, labor and materials, but brings its own productivity residual:

$$\hat{y}_i^t(\text{predicted}) = \exp(\beta_l \bar{l} + \beta_k \bar{k} + \beta_m \bar{m} + \hat{\rho}_i^t) \tag{30}$$

where $\bar{l}$,$\bar{k}$ and $\bar{m}$ are respectively the mean level of salaries, capital and materials over the entire period and $\hat{\rho_i^t}$ is the residual from the production function regression

4 Cost and Production Functions: Duality

Cost functions are closely related to production functions. In fact the cost function is the solution to the following cost minimization problem:

$$\begin{aligned} C(Q,w,r) = \min_{K,L} rK + wL \\ \text{s.t.} \\ Q = \exp(\omega)K^{\beta}L^{\alpha} \end{aligned} \tag{31}$$

Under this assumption of a Cobb-Douglas production function, the Cost function has the following form:

$$C(Q,w,r) = \bar{\omega} + \frac{\alpha}{\alpha+\beta}w_{it} + \frac{\beta}{\alpha+\beta}r_{it} + \frac{1}{\alpha+\beta}q_{it} - \frac{1}{\alpha+\beta}(\bar{\omega} - \omega_{it}) \quad (32)$$

So notice that the original unobserved productivity term ω_{it} is still in the cost equation. Moreover, more productive firms are likely to produce more output and thus q_{it} and ω_{it} are negatively correlated! This would lead us to conclude that firms have *decreasing* returns to to scale since we would see bigger firms have lower unit costs. So you need to use some instruments.

There is some famous work by Nerlove on cost function estimation where he looks the electric power industry and uses the fact that regulators usually tell the utility how much to produce. This is a plausibly exogenous shifter of q_{it}, unless less productive utilities are also allowed to charge higher prices.

5 Reallocation

A principle policy issue are the causes of productivity growth. There is a famous quote by Bob Lucas which states that "Once you start thinking about productivity growth, it is hard to think about much else". Indeed, in the long-run, there is little that is more important than long-run productivity growth to economic outcomes.

An important change in the last 20 to 30 years is the interest in using micro (or establishment level) data to understand the sources of aggregate productivity growth rather than industry or national aggregates. In other words the compositional differences in establishment performance might be important to understand productivity growth. For instance, there is much interest on where the large increases in Chinese productivity are coming from. In particular are they due to the increased importance of the private sector (versus state owned sector) which is composed of more productive establishments, or are all establishments becoming more productive (or perhaps just the ones that export).

Nick Bloom describes the major changes in the micro-view on productivity as:

- High levels of turnover

  About 15% of jobs are destroyed and 20% created in the private sector every year. About 80% of this turnover occurs within the same SIC-4 digit industry. A job is the number of position at an establishments, about 3 times more workers arrive and leave an establishment. (think about turnover at the places where you had summer jobs as a teenager) Davis and Haltiwanger (1991) discuss job creation and destruction in more detail.

- Heterogeneity within industries

  Only about 10% of cross-establishment spread in output, employment, capital and productivity growth is explained by SIC 4-digit controls. Typical gap between 10th and 90th percentiles of productivity within same 4-digit SIC industry is 50%.

- The lumpiness of micro-economic activity

  Many establishments have zero or small levels of investment.

- The importance of reallocation in driving productivity

5.1 Olley-Pakes Decomposition

Olley and Pakes use a decomposition of aggregate productivity to understand where how aggregate productivity growth occurred in the telecommunications industry. Specifically, before ATT was broken up, it's telecommunication equipment manufacturing arm Western Electric had a large share of the sales in the industry, essentially because it was the sole supplier to ATT. However, other telecommunication manufacturing firms may have been more productive, but were prevented from increasing their market share. Table 4 shows this:

Denote industry productivity as $\Omega_t = \sum_i s_{it}\omega_{it}$, where s_{it} is the plant's market share, i.e. $s_{it} = r_{it}/R_t$ where R_t is industry level sales. The OP decomposition comes from the identity:

$$\begin{aligned} \Omega_t &= \sum_i s_{it}\omega_{it} \\ &= \underbrace{\bar{\omega}_t}_{\text{Mean Productivity}} + \underbrace{\sum_i s_{it}(\omega_{it} - \bar{\omega}_t)}_{\text{Covariance of productivity and market share}} \end{aligned}$$

i.e. the average productivity level in the industry plus the covariance of market share and productivity.

In the Olley-Pakes paper itself, instead of using ω the authors use $p_{it} = \exp(\omega_{it})$, i.e. the exponentiated version. Table 5 show their decomposition of productivity.

While there is some evidence of reallocation to more productive producers, it is not clear how significant this effect is.

5.2 Dynamic Decomposition

Another way to decompose productivity growth is to look at covariance based analysis based on *changes* in productivity. To make this analysis easier, let's first focus on the balanced panel case. The dynamic decomposition is given by:

$$
\begin{aligned}
\Delta\Omega &= \sum_i s_{it}\omega_{it} - \sum_i s_{it-1}\omega_{it-1} \\
&= \sum_i s_{it}(\omega_{it} - \omega_{it-1}) \text{ (within)} \\
&+ \sum_i (s_{it} - s_{it-1})\omega_{it} \text{ (between)} \\
&+ \sum_i (s_{it} - s_{it-1})(\omega_{it} - \omega_{it-1}) \text{ (cross)}
\end{aligned}
$$

Now we can also add the role of entrants (henceforth $\mathcal{E}$) and exitors (henceforth χ). Notice, that entrants have $s_{it-1} = 0$ by definition, and entrants have $s_{it} = 0$ by definition. This gives us the expanded decomposition:

$$
\begin{aligned}
\Delta\Omega &= \sum_{i\in I} s_{it}(\omega_{it} - \omega_{it-1}) \text{ (within)} \\
&+ \sum_{i\in I} (s_{it} - s_{it-1})\omega_{it} \text{ (between)} \\
&+ \sum_{i\in I} (s_{it} - s_{it-1})(\omega_{it} - \omega_{it-1}) \text{ (cross)} \\
&+ \sum_{i\in\mathcal{E}} (s_{it})(\omega_{it-1} - \bar{\omega}_t) \text{ (entry)} \\
&- \sum_{i\in\chi} s_{it-1}(\omega_{it-1} - \bar{\omega}_t) \text{ (exit)}
\end{aligned}
$$

Table 6 shows the importance of the within component in manufacturing, taken from Foster, Haltiwanger, and Krizan (2001).

Table 7 from Foster, Haltiwanger, and Krizan (2006) shows that the net entry component is very important component of productivity growth in retail, but not manufacturing.

	Marginal Effect from Probit			
	I	II	III (preferred)	IV
2nd Quintile of Productivity	2.55%	-0.29%	-0.34%	1.63%
	(0.47%)	(0.31%)	(0.31%)	(0.42%)
3rd Quintile of Productivity	1.46%	-1.25%	-1.40%	0.77%
	(0.42%)	(0.29%)	(0.29%)	(0.38%)
4th Quintile of Productivity	-0.28%	-1.77%	-1.74%	-0.59%
	(0.39%)	(0.30%)	(0.30%)	(0.38%)
5th Quintile of Productivity	-1.07%	-2.21%	-2.26%	-1.33%
	(0.44%)	(0.31%)	(0.30%)	(0.37%)
Multi-Unit Status	-4.17%	-4.32%	-4.31%	-4.26%
	(0.26%)	(0.26%)	(0.26%)	(0.26%)
Employment	-0.09%	-0.10%	-0.10%	-0.10%
	(0.01%)	(0.01%)	(0.01%)	(0.01%)
No AR Records			X	
No Hot Imputes		X		
No ASM Years	X			
Pseudo-R2	7.28%	6.95%	6.98%	7.00%
Log Likelihood	-4480.19	-4495.71	-4492.29	-4492.55
Observations	24393	24393	24393	24393
Baseline Exit Probability	3.73%	3.76%	3.75%	3.75%

Table 1: The relationship between productivity and exit is monotonic even after controlling for plant characteristics and imputed data.

Probit on Exit Decision

Baseline Exit Probability = 2.98%

	Marginal Effect	Standard Error
Plant Employees	-0.13%	0.02%
Less than 5 employees	2.38%	0.29%
Multi-Unit Firm	-2.98%	0.20%
1-5 year old plant	1.66%	0.25%
1 year old plant	0.28%	0.28%
Percentile of Productivity	-2.66%	0.47%
Employees*Percentile of Productivity	0.03%	0.01%
Employees*Construction Employment	0.01%	0.00%
Percentile of Productivity*Construction Employment	0.03%	0.03%
Year Fixed Effects	Yes	
Number of Observations	34503	
Pseudo-R2	8.41%	
Log-Likelihood	-5716	

Table 2: A plant at the lowest percentile of productivity has twice the probability of exiting as a plant in the highest percentile of productivity.

TABLE VI

ALTERNATIVE ESTIMATES OF PRODUCTION FUNCTION PARAMETERS[a]

(STANDARD ERRORS IN PARENTHESES)

Sample:	Balanced Panel		Full Sample[c, d]						
								Nonparametric F_ω	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation Procedure	Total	Within	Total	Within	OLS	Only P	Only h	Series	Kernel
Labor	.851	.728	.693	.629	.628			.608	
	(.039)	(.049)	(.019)	(.026)	(.020)			(.027)	
Capital	.173	.067	.304	.150	.219	.355	.339	.342	.355
	(.034)	(.049)	(.018)	(.026)	(.018)	(.02)	(.03)	(.035)	(.058)
Age	.002	−.006	−.0046	−.008	−.001	−.003	.000	−.001	.010
	(.003)	(.016)	(.0026)	(.017)	(.002)	(.002)	(.004)	(.004)	(.013)
Time	.024	.042	.016	.026	.012	.034	.011	.044	.020
	(.006)	(.017)	(.004)	(.017)	(.004)	(.005)	(.01)	(.019)	(.046)
Investment	—	—	—	—	.13	—	—	—	—
					(.01)				
Other Variables	—	—	—	—	—	Powers of P	Powers of h	Full Polynomial in P and h	Kernel in P and h
# Obs.[b]	896	896	2592	2592	2592	1758	1758	1758	1758

[a]The dependent variable in columns (1) to (5) is the log of value added, while in columns (6) to (10), the dependent variable is the log of value added $-b_l * \log(\text{labor})$.

[b]The number of observations in the balanced panels of regressions 1 and 2 are the observations for those plants that have continuous data over the period, with zero investment observations removed. The 2592 observations used in columns (3), (4), and (5) are all observations in the full sample except those with zero investment. Approximately 8% of the full data set had observations with zero investment. Columns (6) to (10) have fewer observations because the sampling procedures for the Annual Survey of Manufactures forced us to drop observations in years 1978, 1983, and the last year, 1987. See note c.

[c]The number of observations in the last four columns decreases to 1758 because we needed lagged values of some of the independent variables in estimation. This rules out using the first observation on each plant and the first year of the rotating five-year panels that make up the Annual Survey of Manufactures. To check that the difference between the estimates in columns (6)–(9) and those in columns (3)–(5) are not due to the sample, we ran the estimating equations in columns (3)–(5) on the 1758 plant sample and got almost identical results.

[d]Consult the text for details of the estimation algorithm for columns (6) to (10).

Table 3: Production Function Estimates from Olley and Pakes (1996).

TABLE II

BELL COMPANY EQUIPMENT PROCUREMENT

(PERCENT PURCHASED FROM WESTERN ELECTRIC)

1982	1983	1984	1985	1986[E]
92.0	80.0	71.8	64.2	57.6

[E]Estimated for 1986.

Source: NTIA (1988, p. 336, and discussion pp. 335–337).

Table 4: ATT Purchases from Western Electric

TABLE XI

DECOMPOSITION OF PRODUCTIVITY[a]
(EQUATION (16))

Year	p_t	$\bar{p}_t$	$\Sigma_i \, \Delta s_{it} \, \Delta p_{it}$	$\rho(p_t, k_t)$
1974	1.00	0.90	0.01	−0.07
1975	0.72	0.66	0.06	−0.11
1976	0.77	0.69	0.07	−0.12
1977	0.75	0.72	0.03	−0.09
1978	0.92	0.80	0.12	−0.05
1979	0.95	0.84	0.12	−0.05
1980	1.12	0.84	0.28	−0.02
1981	1.11	0.76	0.35	0.02
1982	1.08	0.77	0.31	−0.01
1983	0.84	0.76	0.08	−0.07
1984	0.90	0.83	0.07	−0.09
1985	0.99	0.72	0.26	0.02
1986	0.92	0.72	0.20	0.03
1987	0.97	0.66	0.32	0.10

[a]See text for details.

Table 5: Olley-Pakes's OP Decomposition

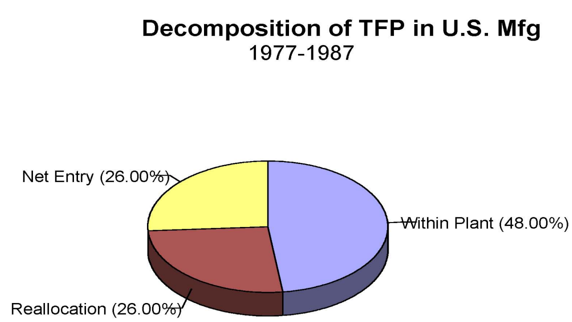

Table 6: In manufacturing, the within component has an important role in productivity growth.

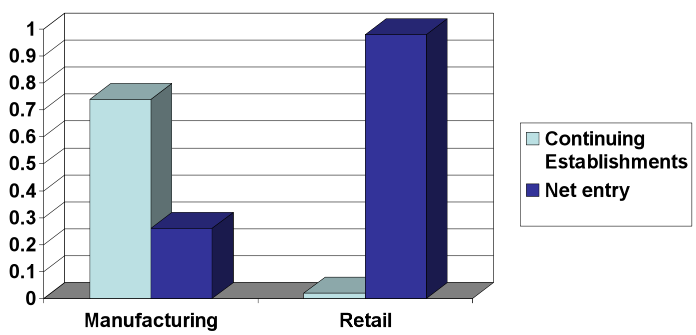

Table 7: In retail, most productivity growth can be traced to the entry of new manufacturing plants.

References

ACKERBERG, D., G. FRAZER, AND K. CAVES (2006): "Structural Estimation of Production Function," *Working Paper UCLA*.

ATHEY, S., AND S. STERN (2002): "The Impact of Information Technology on Emergency Health Care Outcomes," *RAND Journal of Economics*, 33(3), 399–432.

DAVIS, S., AND J. HALTIWANGER (1991): "Gross job creation, gross job destruction and employment reallocation," Discussion paper, National Bureau of Economic Research.

DE LOECKER, J. (2006): "Product Differentiation, Multi-Product Firms and Structural Estimation of Productivity," .

FOSTER, L., J. HALTIWANGER, AND C. KRIZAN (2001): "Aggregate productivity growth. Lessons from microeconomic evidence," .

FOSTER, L., J. HALTIWANGER, AND C. KRIZAN (2006): "Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s," *The Review of Economics and Statistics*, 88(4), 748–758.

LEVINSOHN, J., AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317–341.

OLLEY, G. S., AND A. PAKES (1996): "The dynamics of productivity in the telecommunications equipment industry," *Econometrica*, 64(6), 35.

PAVCNIK, N. (2002): "Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants," *Review of Economic Studies*, 69(1), 245–76.

SYVERSON, C. (2004): "Market Structure and Productivity: A Concrete Example," *Journal of Political Economy*, 112(6), 1181–1222.