

Estimating production functions using  
inputs to control for unobservables

by

James Levinsohn

*National Bureau of Economic Research  
University of Michigan*

and

Amil Petrin

*Graduate School of Business  
University of Chicago*

Current version: April 28, 2000

**Address.** Levinsohn: Department of Economics, University of Michigan, Ann Arbor, MI 48109; Internet: jamesL@umich.edu, WWW: <http://www.econ.lsa.umich.edu/~jamesl>, Petrin: Graduate School of Business, University of Chicago, 1101 E. 58th St., Chicago, IL 60637; Internet: [amil.petrin@gsb.uchicago.edu](mailto:amil.petrin@gsb.uchicago.edu), WWW:<http://gsbwww.uchicago.edu/fac/amil.petrin>

# Estimating production functions using inputs to control for unobservables

James Levinsohn

*National Bureau of Economic Research*

*University of Michigan*

and

Amil Petrin

*Graduate School of Business*

*University of Chicago*

## 1. Introduction

Since at least as early as Marschak and Andrews (1944), applied researchers have worried about the potential correlation between input levels and the unobserved firm-specific shocks in the estimation of production function parameters. The economics underlying this concern are intuitive. Firms that have a large positive productivity shock may respond by using more inputs. To the extent that this is true, Ordinary Least Squares estimates of production functions will yield biased parameter estimates, and, by implication, biased estimates of productivity. The fixed effects solution has the unappealing feature of requiring the transmitted component of the productivity shock to be fixed over time. Instrumental variables is another alternative, but valid instruments need to be correlated with firm-level input choices and orthogonal to the productivity shock. In many cases, there simply are no valid instruments. Olley and Pakes (1996) developed a new approach to addressing this problem— one which did not require instruments, and we take their approach (reviewed below) as our starting point.

We introduce a new method for conditioning out serially correlated unobserved shocks to the production technology by building ideas first developed in Olley and Pakes (1996). Olley and Pakes show how to use investment to control for correlation between input levels and the unobserved firm-specific productivity process. We prove that like investment, intermediate inputs can also solve this simultaneity problem. We highlight three potential advantages to using an intermediate inputs approach relative to investment.

---

We would like to thank seminar participants at UC Berkeley, Toronto, Yale, University, Harvard University, University of Chicago, and NBER for helpful suggestions on earlier work on this project. Susanto Basu, Peter Klenow, and Ariel Pakes provided especially helpful suggestions. Wendy Petropoulos provided splendid research assistance and many helpful ideas. We are grateful to the Russell Sage Foundation and the Centel Foundation/Robert P. Ruess Faculty Research Fund at the GSB, the University of Chicago for support.

The first advantage is that intermediate inputs will generally respond to the *entire* productivity term, while investment may respond only to the “news” in the unobserved term. This can happen for two reasons. If the capital input has already adjusted to the “forecastable” component of the productivity process, the investment proxy will only account for the “non-forecastable” component of productivity. Also, productivity may be characterized by two components, a serially correlated component to which investment responds, and a separate firm-time shock that is independent over time, to which investment will not respond, but to which the choice of variable factors will respond. In both cases, the investment proxy will only account for the part of the productivity term to which capital responds, and some correlation between the regressors and the error term will still remain. Intermediate inputs, on the other hand, will generally respond to the entire shock, and will hence restore consistency.

A second advantage is that intermediate inputs provide a simpler link between the estimation strategy and the economic theory, primarily because intermediate inputs are not typically state variables. We develop this simpler link and discuss the conditions that must hold if intermediate inputs are to be a valid proxy for the productivity shock.

The third advantage is strictly data-driven. It turns out that the investment proxy is only valid for firms reporting non-zero investment. (This is due to an invertibility condition described below.) Pronounced adjustment costs, which do not invalidate the use of investment as a proxy, are the likely reason that over one-half of our sample reports zero investment. We are concerned about the possible truncation bias the exclusion of “zero investment” firms might introduce. Using intermediate input proxies avoids the potential truncation of a large number of firms in industries with pronounced adjustment costs of capital. This is because, in general, firms always report positive use intermediate inputs like electricity or materials.

In an effort to ascertain whether intermediate inputs address the simultaneity in practice, we compare estimates across OLS, the Olley-Pakes investment proxy estimator, and our intermediate input proxy estimator. Using data on eight different Chilean manufacturing industries, we demonstrate that the intermediate inputs proxy is important for obtaining consistent and unbiased estimates of the parameters of the production function.

The remainder of the paper is organized as follows. Section 2 provides a very brief review of the simultaneity problem. In Section 3, we introduce our intermediate input proxy, and develop the conditions under which it will be a valid estimator. Section 4 describes our data, and Section 5 includes the details of the estimation approach. In section 6 we present our results, while Section 7 concludes. An appendix includes a “recipe” for our estimation routine.

## 2. Estimation in the Presence of Simultaneity

In this section, we briefly review the simultaneity problem described by Marschak and Andrews (1944), including its consequences for some commonly used estimators for production function parameters.<sup>1</sup> We then describe the Olley and Pakes (1996) simultaneity correction. Finally, we discuss when the investment proxy might fail to adequately address the simultaneity problem.

### *The Simultaneity Problem*

We begin our discussion of the simultaneity problem by placing it in the context of a profit-maximizing firm. In a discrete time model, the expected total discounted profits for firm  $i$  can be written as

$$\Pi_1(Y_{i1}, X_{i1}, \epsilon_{i1}; \theta) + E\left[\sum_{t=2}^{\infty} \beta^{t-1} \Pi_t(Y_{it}, X_{it}, \epsilon_{it}; \theta)\right], \quad (1)$$

where  $\Pi_t(\cdot)$  is the profit function for period  $t$ , and  $\beta$  is the discount rate.  $Y_{it}$  is a measure of output and  $X_{it}$  is a vector of inputs for period  $t$ .  $X_{it}$  includes inputs that are easily adjusted (e.g. materials or labor, in many cases.)  $X_{it}$  also includes inputs with adjustments costs, that is, whose stock evolves over time in response to both today's *and* future beliefs (e.g. capital.) A sequence of errors,  $\{\epsilon_{it}\}_{t=1}^{\infty}$ , indexed both by firm and time, accounts for differences between the model's predictions and observed outcomes. These shocks will include components that are transmitted over time (and hence taken into account by managers) and those that are not.

A simultaneity problem arises when there is contemporaneous correlation both within firm  $i$  and across time  $t$  between  $\epsilon_{it}$  and  $X_{it}$  in the firm-specific sequences  $\{X_{it}, \epsilon_{it}\}_{t=1}^{\infty}$ . Given the natural dependence of the firm's discounted future profits on both  $X_{it}$  and  $\epsilon_{it}$  in (1), this correlation seems plausible. For over 50 years, the potential correlation has been pointed out, and applied researchers have spent much effort addressing the econometric problem it confers.

Marschak and Andrews (1944) argue that the problem arises when variable input demands are correlated with unobserved shocks. They suggest that this simultaneity can be particularly acute for inputs that adjust most rapidly to the shocks. There is also reason to believe that firms with better sequences of shocks will, over time, respond to these "good" realizations by investing and accumulating assets that are costly to adjust rapidly. Alternatively, simultaneity may occur in the time dimension when input decisions are based on serially correlated errors and there are costs to making large immediate adjustments to inputs.

---

<sup>1</sup> Some of this discussion parallels that found in Griliches and Mairesse (1995) See that paper for a more extensive discussion of the issues discussed in this section.

Econometrically, with simultaneity it is generally impossible to sign the biases of the production function coefficients when there are many inputs, all of which may be (to differing degrees) correlated with the error. However, some intuition can be derived from an analysis of the OLS estimates for a two input production function, with output  $y_{it}$ , one freely variable input  $l_{it}$  (call it labor) and one quasi-fixed input  $k_{it}$  (call it capital):

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \epsilon_{it}.$$

One can show that the OLS estimates for the inputs are

$$\hat{\beta}_l = \frac{m_{y,l}m_{k,k} - m_{y,k}m_{l,k}}{m_{l,l}m_{k,k} - m_{l,k}^2},$$

and, symmetrically,

$$\hat{\beta}_k = \frac{m_{y,k}m_{l,l} - m_{y,l}m_{l,k}}{m_{l,l}m_{k,k} - m_{l,k}^2},$$

where  $m_{a,b}$  denotes the sample covariance between  $a$  and  $b$ . If capital is not correlated with labor (which is not true in our data across firms) and labor is positively correlated with productivity, then  $\hat{\beta}_l$  is biased up but  $\hat{\beta}_k$  unbiased. In the case where capital and labor are positively correlated, the positive bias to the labor coefficient gets transmitted to the capital coefficient, which becomes biased down. Finally, in the case where capital and labor are positively correlated, and both capital and labor are also correlated with the productivity shock, but labor's correlation is significantly stronger, the formulas above suggest that  $\hat{\beta}_l$  will again tend to be overestimated and  $\hat{\beta}_k$  will tend to be underestimated. Given what we know about our data, our prior is that this last case is probably most likely to be true.

OLS and the within estimators are the most common methods for estimating production function parameters. However, they impose very strict restrictions on the sequences  $\{X_{it}, \epsilon_{it}\}_{t=1}^{\infty}$  by ruling out, *a priori*, important forms of firm-level heterogeneity across establishments and over time. OLS assumes that  $\epsilon_{it}$  is uncorrelated with input choices across both firms and time. Within estimators protect against firm-specific effects, but assume that there is no transmitted component to the firm-level error. Within estimators also dispense with all between variation, which can be particularly critical for obtaining precise estimates of output elasticities associated with state variables. In this sense, within estimators offer more protection against firm-specific effects than OLS, but they can exacerbate other problems (like measurement error in capital) by dispensing with much of the “signal”. This has led people to the instrumental variables approach.

The instrumental variable estimator starts from the premise that  $X_{it}$  and  $\epsilon_{it}$  may be correlated. The approach is to find variables that are correlated with  $X_{it}$  but uncorrelated  $\epsilon_{it}$ .<sup>2</sup> Unfortunately, it can be difficult to locate instruments for inputs.<sup>3</sup> Input prices that vary across firms and/or over time might qualify. Lagged values of inputs will not generally be valid instruments because chosen input levels may depend upon past values of the (potentially correlated) shock. Frequently, instrumental variables suffers from the same drawback as that of the within estimator; valid instruments are usually weak instruments - that's generally what makes the exclusion restriction believable - and weak instruments significantly weaken the signal, exacerbating other imperfections in the data.

### *The Investment Proxy*

Olley and Pakes (1996) suggest a novel approach to addressing the simultaneity problem as it arises in the estimation of production functions. They abandon the traditional “solutions” of instrumental variables or within estimators. Instead, they include in the estimation equation a proxy for the productivity term with motivation for the chosen proxy derived from a structural model of the optimizing firm. The proxy controls for the part of the error correlated with inputs. That is, it dispenses with all variation in output and inputs that is related to the productivity term; hence, identification relies only on variation in output and inputs unrelated to productivity.

We simplify (slightly) their model, writing the production function in logs as

$$y_t = \beta_0 + l_t\beta_l + \beta_k k_t + \omega_t + \eta_t, \tag{2}$$

where lower case letters will always denote log-levels (the firm index is suppressed in what follows.)

---

<sup>2</sup> It is also possible to directly specify the parametric process that the productivity term follows. However, even if we are willing to characterize the dynamic sequence  $\{X_{it}, \epsilon_{it}\}_{i=1}^{\infty}$  as a parametric process and want only to estimate the parameters of this process, we still have a significant problem. By itself, knowledge of the process (up to the parameters) is not enough to control for the simultaneity between  $\epsilon_{it}$  and  $X_{it}$  over time because the process  $\{X_{it}, \epsilon_{it}\}_{i=1}^{\infty}$  follows a path that depends upon its starting values  $(X_{i1}, \epsilon_{i1})$ . This is an initial conditions problem (see Heckman (1981) and Pakes (1996)), where estimation of parameters for a stochastic process that depends upon time-ordered outcomes is impossible unless the process is “initialized.” One solution is to initialize the observed process by assuming the history is exogenous, i.e. that  $\{X_{it}, \epsilon_{it}\}_{i=1}^{T-1}$  is independent of  $\{X_{it}, \epsilon_{it}\}_{i=T}^{\infty}$ , where  $T$  is the first date a firm is observed. Another solution is to estimate a fixed entry-time effect for each firm. However, this solution produces consistent estimates only as the ratio (observations/firm) gets large. This estimator is not consistent as just the number of firms increases because each extra firm brings with it an extra fixed effect parameter to estimate. A third solution splits the sample into two parts, the first part of which is used to estimate starting values. Roberts and Tybout (1997) take this approach with a panel from Colombia that is 10 years in length. Using the first 3 years they estimate starting values for continuing firms, and then initialize their assumed stochastic process accordingly. The second half of the data set is then used to look for determinants of firms’ decisions to enter the export market in Columbia.

<sup>3</sup> This difficulty is especially acute when using plant-level data. In macroeconomics, there is a practice of estimating what are called production functions using industry level data and in this case, there *are* plausible instruments. The problem here is that industries don't actually produce anything, plants do.

Inputs are divided into a freely variable one ( $l_t$ ), and the state variable capital ( $k_t$ ).<sup>4</sup> Additionally, (the growth rate of)  $\epsilon_t$  is assumed to be additively separable in two components, a transmitted component ( $\omega_t$ ), and an i.i.d. component ( $\eta_t$ ). The key difference between  $\omega_t$  and  $\eta_t$  is that the former is a state variable, and hence impacts the firm’s decision rules, while the latter has no impact on the firm’s decisions.

In a perfectly competitive environment where input and output prices are common across firms, the capital control investment can be written as just a function of the two state variables,  $k_t$  and  $\omega_t$ , or

$$i_t = i_t(\omega_t, k_t).$$

Pakes (1996) proves (under certain conditions) that optimizing firms *choosing to invest* have investment functions that are strictly increasing in the unobserved productivity shock.<sup>5</sup> Hence,  $i_t(\omega_t, k_t)$  can be inverted to yield  $\omega_t$  as a function of investment and capital, or  $\omega_t = (\omega_t, k_t)$ . Given this “monotonicity” condition, one can rewrite (2) as

$$y_t = l_t\beta_l + \phi_t(i_t, k_t) + \eta_t, \tag{3}$$

where

$$\phi_t(i_t, k_t) = \beta_0 + \beta_k k_t + \omega_t(i_t, k_t).$$

A first-stage semi-parametric estimator (non-parametric in  $\phi_t$ ) can then be used to obtain consistent estimates of the coefficients on the freely variable inputs.<sup>6</sup> Again, the idea is to identify the variable input coefficients using *only* the “qualified variation,” that is, the variation unrelated to  $i_t$  and  $k_t$ . Olley and Pakes achieve this goal by including a polynomial in  $i_t$  and  $k_t$  in the regression of output on the variable inputs; as long as there are “enough” terms in the expansion, this approach will achieve the aforementioned goal.

A different and perhaps slightly more transparent way to see that this goal is achieved is to follow the exposition in Robinson (1988). His discussion has the added advantage of illustrating

---

<sup>4</sup> For simplicity, we assume (as they do) that capital is the only state variable over which the firm has control. In principal, Olley-Pakes extends to frameworks with many state variables.

<sup>5</sup> If the distribution of next period’s productivity is stochastically increasing in this period’s productivity, the economic story that makes investment a valid proxy is straightforward; a firm that realizes large productivity this period will invest more than an identical firm with smaller productivity this period because the more productive firm anticipates doing better than the less productive firm both in the current period and in future periods.

<sup>6</sup> We will always use  $\phi_t(\cdot)$  when discussing the non-parametric part of this first stage; it’s arguments will change, but it will always include capital and the proxy variable. More generally,  $\phi_t(\cdot)$  will always have as arguments all of the endogenous state variables and the proxy variable.

how one might implement alternative non-parametric estimators (in addition to the polynomial approximation used by Olley and Pakes) to check the robustness of the results to different approaches. Robinson (1988) proceeds by constructing the expectation of equation (3) conditional on  $i_t$  and  $k_t$ . This expectation is given by

$$E[y_t|i_t, k_t] = E[l_t|i_t, k_t]\beta_l + \phi_t(i_t, k_t) \quad (4)$$

because: i)  $\eta_t$  is mean independent of  $i_t$  and  $k_t$ ; and ii)  $E[\phi_t(i_t, k_t)|i_t, k_t] = \phi_t(i_t, k_t)$ . Subtracting equation (4) from equation (3) (i.e. netting out the variation in output and inputs associated with variation in  $i_t$  and  $k_t$ ) yields

$$y_t - E[y_t|i_t, k_t] = (l_t - E[l_t|i_t, k_t])\beta_l + \eta_t. \quad (5)$$

Since the  $\phi_t$  differences out, the error in this new equation ( $\eta_t$  instead of  $\omega_t + \eta_t$  from (2)) is mean independent of the transformed regressor  $l_t - E[l_t|i_t, k_t]$ . Hence no-intercept OLS can be used to obtain consistent estimates of the coefficient on the freely variable input by regressing  $y_t - E[y_t|i_t, k_t]$  on  $l_t - E[l_t|i_t, k_t]$ .

A second stage is required to obtain an estimate of the capital coefficient, and we briefly discuss this now. Olley and Pakes assume  $\omega_t$  follows a first-order Markov process, so the expected value (or best guess) of  $\omega_t$  conditional on knowing  $\omega_{t-1}$  is the same as the expected value given  $\omega_{t-1}$ ,  $\omega_{t-2}$ , etc..<sup>7</sup> One can decompose  $\omega_t$  into two components,

$$\omega_t = E[\omega_t|\omega_{t-1}] + \xi_t,$$

the expected value of  $\omega_t$  given  $\omega_{t-1}$ , and the mean zero innovation or “surprise” in  $\omega_t$ , or  $\xi_t$ . While the first stage of the estimation routine allows the variable inputs to respond to both  $\xi_t$  and  $E[\omega_t|\omega_{t-1}]$ , identification of  $\beta_k$  will obtain from the assumption that capital slowly adjusts to  $\omega_t$ ; in period  $t$ , capital is assumed to only respond to  $E[\omega_t|\omega_{t-1}]$ , and thus is mean independent of  $\xi_t$ .

The second stage begins by netting from output the contribution of  $l_t$ , or  $l_t\beta_l$ , obtaining a new dependent variable  $y_t^*$ , or

$$y_t^* = y_t - l_t\beta_l = \beta_0 + \beta_k k_t + \omega_t + \eta_t. \quad (6)$$

If one collects the intercept and  $E[\omega_t|\omega_{t-1}]$  into one function  $g(\omega_{t-1})$ , or

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

---

<sup>7</sup> This approach immediately generalizes to higher-order Markov processes.



then (6) becomes

$$y_t^* = \beta_k k_t + g(\omega_{t-1}) + \eta_t^*, \quad (7)$$

where  $\eta_t^* = \xi_t + \eta_t$ .<sup>8</sup> (7) closely resembles (3), so the same approach and intuition for estimating (3) is applicable to (7); only variation in  $y_t^*$  unrelated to  $g(\omega_{t-1})$  can be used to estimate  $\beta_k$ . And since a by-product of the first stage is an estimator for  $\omega_{t-1}$ , estimation conditional on  $g(\omega_{t-1})$  is possible. (In practice, estimation of  $\beta_k$  is more involved than OLS on transformed variables since  $g(\omega_{t-1})$  will itself depend on  $\beta_k$ .) Again, identification obtains from assuming that capital is slow to adjust to the innovations in the transmitted component of the error, so (conditional on  $\omega_{t-1}$ ),  $k_t$  is mean independent of the new error  $\eta_t^*$ .

When this approach works, it has a number of advantages over OLS, within, and traditional instrumental variable estimators. Griliches and Mairesse (1995) document these advantages (we change the variable references so as to be consistent with our notation):

The major innovation of Olley and Pakes is to bring in a new equation, the investment equation, as a proxy for  $\omega$ , the unobserved transmitted component of  $\epsilon$ . Trying to proxy for the unobserved  $\omega$  (if it can be done right) has several advantages over the usual within estimators (or the more general Chamberlain and GMM type estimators): it does not assume that  $\omega$  reduces to a “fixed” (over time) firm effect; it leaves more identifying variance in  $x$  and  $k$ , and hence is a less costly solution to the omitted variable and/or simultaneity problem; and it should also be substantively more informative.

However, this approach will not always work, and we now turn to some examples where the simultaneity problem is not solved by the investment proxy.

#### *When The Investment Proxy Might Fail*

Our motivation for introducing the intermediate input proxy arises primarily from the concern that investment may not respond to the entire transmitted shock. The problem arises principally because investment is the control on a *state* variable, and a state variable is by definition costly to adjust. These costs of adjustment take different forms. Our concern is that the adjustment cost may affect the responsiveness of investment to the transmitted shock in a manner which would violate the monotonicity condition. We describe two such scenarios.

First, the transmitted component of the error (or productivity, if you like) may have a forecastable component and a non-forecastable component. If part of the transmitted shock is predictable, the adjustment process will accommodate itself in an optimal manner to this predictable part. For example, better firms may know that they are better, and may adjust their capital

---

<sup>8</sup>  $\beta_0$  and the intercept from  $E[\omega_t | \omega_{t-1}]$  are not separately identified without some additional assumption.

stocks accordingly. If capital has already adjusted to this component of the shock, investment will only respond to the “news” in the shock. Hence, the proxy will only control for this “news”. Of course, the freely variable inputs respond to *both* the news and the forecastable component, so the simultaneity problem still exists.

A second scenario might also lead to the failure of the investment proxy. If there is an i.i.d. (non-transmitted) component to each period’s shock that enhances productivity, it will not change expectations about the future. However, it will affect input levels for the freely variable factors, inducing a correlation between this component of the error and variable inputs. Since the investment proxy fails to capture this component of the “productivity” shock, the parameter estimates will again not be consistent. We next turn to a proxy that will be robust to these and other more general processes.

### 3. Intermediate Inputs as Proxies

In this section we introduce the intermediate input proxy. We develop the conditions under which an intermediate input can solve the simultaneity problem. We show that an intermediate input, when valid as a proxy, will be robust to the problems associated with the investment proxy discussed above. We also find that establishing the validity of an intermediate input proxy can be significantly easier than establishing this fact for investment.

Two key conditions must be met if intermediate inputs are to be a valid proxy. The first condition, *Condition 1*, is a monotonicity condition; conditional on capital, intermediate input use must increase in  $\omega$ . *Condition 2* is that the market environment is competitive. We also discuss a *Condition 3*– separability of the production technology in the intermediate input used as the proxy. While this condition need not necessarily hold, we develop the idea because it is often assumed to hold in practice, e.g. it must hold *whenever* a value-added production function is estimated. In fact, estimating a value-added production function requires stronger conditions than our *Condition 3*; separability must hold for *all* inputs excluded from the value-added production function. Since this stronger condition is frequently assumed to hold, we show how it allows for an alternative solution to the simultaneity problem if the revenue share of the intermediate input is observed.<sup>9</sup>

In the next section, we discuss each of these conditions in turn. We proceed by positing the first two conditions and deriving the resulting estimating equations. We then show how the third condition can be useful. Finally, we return to examine in more detail each of the three conditions and the economics underlying the estimating equations.

---

<sup>9</sup> This is almost always observed in current firm-level data sets.

### *Using Intermediate Input Proxies*

Here we provide an overview of the estimation approach, describing where each of the different conditions play their role. Before we begin, we note that our estimation approach applies to quite general production technologies, including those with many freely variable inputs and many state variables, and with higher order and interaction terms. However, for ease of exposition and because nothing from a methodological standpoint is gained from using a more general representation of the technology, we stick to a first-order Taylor series approximation with three inputs.<sup>10</sup>

As before, we write the freely variable input, labor, as  $l$ , the state variable capital as  $k$ , the “productivity” shock as  $\omega$ , and the shock to which all inputs do not respond as  $\eta$ . We introduce the second freely variable input,  $\iota$ , which we call the intermediate input. Writing the log of output as a function of the log of inputs and the shocks we have

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \beta_\iota \iota_t + \omega_t + \eta_t. \quad (8)$$

We start with the intermediate input’s demand function, which we write as

$$\iota_t = \iota_t(\omega_t, k_t).$$

By writing the intermediate input demand function as only a function  $\omega_t$  and  $k_t$ , we have implicitly invoked *Condition 2*, or perfect competition, where input and output prices are identical across firms. If firms faced different input prices or charged different output prices, the function would also have to be indexed by these prices. By indexing the input demand function by  $t$ , we do however allow these prices to change over time.

Next, we invoke *Condition 1*, or the monotonicity condition, which must hold for the intermediate input  $\iota$  so we can invert the intermediate input demand function and write

$$\omega_t = \omega_t(\iota_t, k_t).$$

Substituting the proxy into (8) yields:

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \beta_\iota \iota_t + \omega_t(\iota_t, k_t) + \eta_t. \quad (9)$$

Expressing (9) in a manner similar to (3), we get

$$y_t = \beta_l l_t + \phi_t(\iota_t, k_t) + \eta_t, \quad (10)$$

---

<sup>10</sup> Exact details of the estimation routine for a trans-log production function with three variable inputs and two state variables are available from the authors by request.

where

$$\phi_t(l_t, k_t) = \beta_0 + \beta_k k_t + \beta_l l_t + \omega_t(l_t, k_t).$$

The coefficient on the variable input can then be estimated using the semi-parametric estimator of Robinson (1988) described earlier in the *Investment Proxy* section, where no-intercept OLS is used to estimate the equation

$$y_t - E[y_t | l_t, k_t] = (l_t - E[l_t | l_t, k_t])\beta_l + \eta_t. \quad (11)$$

Comparing (11) with (5) illustrates exactly how our approach is robust to changes in productivity to which the investment proxy may not respond; the only “qualified” variation in output and inputs that is used to identify the coefficient on the variable input is variation *unrelated* to movements in the intermediate input and capital (instead of variation unrelated to movements in investment and capital.)

In summary, the first stage identifies the coefficients on the variable inputs (except the coefficient on the proxy input.) Hence, if the economic question at hand requires no more than an estimate of the elasticity of output with respect to any (or all) of these inputs, we would be done. However, in general one also wants estimates of  $\beta_k$  and  $\beta_l$ , either as parameters of interest themselves, or because one is estimating productivity and/or returns to scale. A second stage of estimation is necessary to obtain these estimates.

The theory for the second stage begins with one netting from output the contribution of  $l_t$  to output, obtaining a new dependent variable  $y_t^*$ , or

$$y_t^* = y_t - l_t \beta_l = \beta_0 + \beta_k k_t + \beta_l l_t + \omega_t + \eta_t. \quad (12)$$

As described earlier, one can decompose  $\omega_t$  into two components,  $\omega_t = E[\omega_t | \omega_{t-1}] + \xi_t$ , the expected value given  $\omega_{t-1}$  and the “surprise” in  $\omega_t$ , or  $\xi_t$ .<sup>11</sup> If one then collects the intercept and  $E[\omega_t | \omega_{t-1}]$  into one function  $g(\omega_{t-1})$ , or

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

then (6) becomes

$$y_t^* = \beta_k k_t + \beta_l l_t + g(\omega_{t-1}) + \eta_t^*, \quad (13)$$

where  $\eta_t^* = \xi_t + \eta_t$ . (13) closely resembles (7), except for the inclusion of the term related to the intermediate input. While the same approach and intuition for estimating (7) is again applicable

---

<sup>11</sup> The “surprise” interpretation of  $\xi_t$  follows from the assumption that  $\omega_t$  follows a first-order Markov process.

for (13) when considering capital, it does *not* hold for the intermediate input; the intermediate input is correlated with  $\eta_t^*$  because it responds to the innovation  $\xi_t$  in  $\omega_t$ . Thus neither coefficient is identified without some further restriction.

We suggest two ways to identify  $\beta_k$  and  $\beta_l$ . The first approach, which we call the unrestricted approach, uses the fact that  $\iota_{t-1}$  is uncorrelated with  $\eta_t^* = \xi_t + \eta_t$ , and hence qualifies as a valid instrument. That is,  $\iota_t$  and  $\iota_{t-1}$  may be correlated within a firm, but last period's intermediate input use will be orthogonal to the unexpected part of this period's productivity shock. When this observation is combined with the assumption that capital does not respond to the innovation in productivity, we then have two moment conditions that we can use to identify the two coefficients  $\beta_k$  and  $\beta_l$ .

A second approach is available if the revenue share of an input is observable and the production technology is separable in this input.<sup>12</sup> Under perfect competition, cost minimization implies that firms set the revenue share of the intermediate input equal to the elasticity of output with respect to that input, or

$$\beta_l = s_l. \tag{14}$$

Hence *Condition 2* and *Condition 3* together allow the intermediate input contribution to output to be expressed as  $\beta_l \iota_t = s_l \iota_t$ . One can then net out from equation (13) an estimate of  $\beta_l \iota_t$  and proceed with estimation in exactly the same way as with equation (7). We refer to this second approach as the restricted approach. We now describe in more detail each of these three conditions.

### *The Monotonicity Condition*

We require a monotonicity condition in order to be able to invert the intermediate demand function and express the transmitted productivity shock as a function of the intermediate input and capital.

The monotonicity condition for intermediate inputs is identical to that for investment; conditional on capital, profit maximizing behavior must lead more productive firms to use more intermediate inputs. The story we imagine behind valid intermediate input proxies is straightforward. If an increase in productivity leads to higher marginal products for capital, and/or labor, an optimizing firm facing constant output and input prices will want to produce more output, doing so until the marginal products fall to the point where they equal the (normalized) input prices. In the process, the firm will increase the use of inputs. Hence, positive shocks lead to increases in input

---

<sup>12</sup> Revenue share is defined as the amount spent on the intermediate input divided by the total firm revenue from output.

use. Finally, if larger shocks lead to bigger increases in the marginal products (relative to smaller shocks), then, conditional on capital, more “productive” firms will demand more intermediate inputs relative the less productive firms. This result is the monotonicity result. In summary, as long as (conditional on capital) more productive firms find it more profitable to produce more output than less productive firms, the more productive firms will use more intermediate inputs to achieve this higher level of output, and the monotonicity condition will hold.

We now turn to the condition that must hold on the production technology for the monotonicity result to obtain when input and output prices are fixed (that is, when firms can supply as much output as they wish at a given price.)<sup>13</sup> Let

$$Y = f(L, K, \iota, \omega)$$

represent the production technology, where output is determined by labor (L), capital (K), intermediate inputs ( $\iota$ ), and productivity ( $\omega$ ). Aside from some regularity conditions on  $f(\cdot)$ , conditional on  $K$ , intermediate input use will be increasing in  $\omega$  if

$$f_{\iota L} f_{L\omega} > f_{LL} f_{\iota\omega} \tag{15}$$

holds at all input combinations relevant for production (where  $f_{LL}$  is the second derivative of  $f(\cdot)$  with respect to  $l$ , etc.) The result obtains because

$$\text{sign}\left(\frac{d\iota}{d\omega}\right) = \text{sign}(f_{\iota L} f_{L\omega} - f_{LL} f_{\iota\omega});$$

when (15) holds, small increases in  $\omega$  lead to small increases in intermediate input use. The fundamental theorem of calculus then establishes the result for large discrete changes, or

$$\iota(\omega_2, K_1) > \iota(\omega_1, K_1)$$

whenever  $\omega_2 > \omega_1$ .

Economic theory is useful for determining when (15) is likely to hold. Optimizing behavior implies the marginal product of labor declines as labor increases, so  $f_{LL}$  will be negative. Additionally, if increases in productivity always weakly increase the marginal product of inputs, then  $f_{L\omega}$  and  $f_{\iota\omega}$  are non-negative (one must be positive.) The result is then driven by the cross-partial of output with respect to the intermediate input and labor. If the marginal product of the intermediate input weakly increases as labor use increases (i.e.  $f_{\iota L} \geq 0$ ), then the monotonicity result

---

<sup>13</sup> Appendix A contains the full derivation.

holds. However, if the marginal product of the intermediate input falls rapidly with increases in labor, then  $(f_{lL}f_{L\omega} - f_{LL}f_{l\omega})$  may be negative; in this case, the monotonicity result could (but does not have to) fail.

Before turning to the estimation, we point out another potential advantage of using an intermediate input proxy; the condition on the production technology is easy to verify, and is satisfied by many technologies used by economists. This result contrasts with the proof that investment is monotonic in productivity (see Pakes (1996).) If one wishes to use a model that differs, even slightly, from that of Olley and Pakes, it becomes necessary to re-investigate the appropriateness of the investment proxy. The fact that investment is a control on a state variable forces one to consider the firm's dynamic problem when showing the monotonicity condition for investment. In his original monotonicity proof, Pakes (1996) demonstrates that this can be tricky business.

### *The Perfect Competition Condition*

The need for competition relates to the fact that input and output prices are not observed at the plant level. However, if the environment is competitive, then we know that plants face the same input and output prices, and this information can be used to address a number of estimation problems, which we now discuss.

First, as is true with almost all available plant-level data, we do not observe output. Instead, we observe total revenue, and we deflate this total revenue by an industry price deflator to account for variation in prices over time. This price deflator is common across plants, and thus plants must face a common price for this deflation to be technically correct.

Second, when we write the intermediate input demand function  $\iota_t(\omega_t, k_t)$ , we implicitly assume that firms face the same output and input prices; these prices affect the input demand functions but we don't index the function by them. The competitive environment ensures these prices are common across firms, thus allowing us to estimate this function with just the two state variables as explanatory variables. Of course, if input and output prices were observed (and they differed across firms), we would just index this function with these prices. As noted before, we do index this function by time, which allows input and output price ratios to differ in an arbitrary way across the time periods.

A third reason perfect competition is useful relates to the monotonicity condition. It is straightforward to show monotonicity under perfect competition, because when firms face common input and output prices, they will generally want to produce more output if they are more productive, and hence will use more intermediate inputs. If the underlying market structure is not perfectly

competitive, it is no longer obvious that a firm that enjoys a larger productivity shock will necessarily produce more (and hence drive down marginal productivities.) This is because the oligopolistic firm realizes that by producing more, price too is driven down. In many, but not all, market structures, the monotonicity condition may generally hold. (It is true, for example, in the simple homogeneous products Cournot case with linear demands.) However, it is clear that one must make sure it holds for underlying imperfectly competitive market structure under consideration. While addressing these issues and relaxing the perfect competition assumption is the subject of ongoing work, the point here is simply to highlight why perfect competition is assumed.

### *The Separability Condition*

We now discuss *Condition 3*, which requires the production function to be weakly separable in the particular input that is used as a proxy. We show that if the input used as a proxy is separable and the market environment is competitive, it is possible to net from the output variable the contribution of the input. We remind the reader we do not need this condition; we simply provide it as a possible alternative because conditions stronger than this condition are so frequently invoked, e.g. whenever a value added production function is estimated, all of the inputs that are subtracted from output before estimation proceeds must technically enter the production technology in a separable way; otherwise, they cannot simply be subtracted out.<sup>14</sup>

Here we follow the lead of Solow (1957), and Griliches and Ringstad (1971), who both suggest (for different reasons) netting the contribution of inputs from output. In particular, Griliches and Ringstad (1971), who are primarily interested in estimating the capital coefficient, address the simultaneity problem by netting from output the contribution of labor. This contribution is computed by: i) assuming the output function is separable in labor; and ii) invoking profit maximization to obtain the labor input contribution to output - the product of the revenue share of the variable input with the (log) level of this same input. Later we present results that do not directly estimate the coefficient on the intermediate input in the second stage. Instead, we use the revenue share times the intermediate input level as a measure of the intermediate input's contribution to output (which is then net from output when estimating the capital coefficient on our second stage.) In our data, we find that this restricted approach yields very similar results to the unrestricted approach that only uses the first two conditions.

---

<sup>14</sup> See Bruno (1978) and Basu and Fernald (1995) for careful discussions of estimation using value added production functions.



We would like to note again that our general approach does not require that the production technology be separable in *all* of the intermediate inputs (the standard assumption when a value added function in labor and capital is estimated.) All we need is for one input to be separable. For example, if both raw materials and electricity satisfied the monotonicity condition, we could choose the one that we think is most likely to satisfy the separability condition if we want to use *Condition 3*. We now turn to the data.

#### 4. Data

In order to implement the intermediate input proxy, we need data. We choose as our test ground an eight year panel from Chile that has been used elsewhere.<sup>15</sup> This Chilean data has many firm-level variables, is not censored for entry and exit, and has a reasonable time-series dimension to it. We hope these characteristics are representative of many firm-level panels.<sup>16</sup>

We first provide a description of the variables used in the analysis. We then present some summary statistics on each of the eight industries we examine. Next, we report the number of observations which would be lost using any particular input as a proxy, and we explain why we choose electricity as our primary proxy.

The data set is comprised of plant-level data of 6665 plants in Chile from 1979 to 1986. The data are a manufacturing census covering all plants with at least ten employees. The data were originally provided by Chile's Instituto Nacional de Estadística (INE). A very detailed description of how the eight longitudinal samples were combined into a panel is found in Lui (1991). There is information tracking plants over time, including plants that enter over the course of the sample period (births) as well as plants that exit (deaths.)<sup>17</sup>

In an attempt to keep the analysis manageable, we focus on eight of the largest industries (excluding petroleum and refining.) We work with industries at the 3-digit level. The industries (along with their ISIC codes) are Metals (381), Textiles (321), Food Products (311), Beverages (313), Other Chemicals (352), Wood Products (331), and Printing and Publishing (342), and

---

<sup>15</sup> See, for example, Lui (1993), Lui and Tybout (1996), Tybout, de Melo, and Corbo (1991), Levinsohn (1998), and Pavcnik (1999).

<sup>16</sup> Unlike the LRD used by Olley and Pakes, the Chilean data is easily accessed so others can readily replicate and extend our analysis.

<sup>17</sup> Due to the way that the data are reported, we treat plants as firms, although there are certainly multi-plant firms in the sample. We will not capture the extent to which multi-plant firms experience scale or scope economies due to their multi-plant nature. Neither are we able to investigate whether "entry" is a new firm, a new plant from an existing firm, or simply diversification of an existing plant or firm as discussed in Dunne, Roberts, and Samuelson (1988).

Apparel (322). The data are observed annually and they include an index of output, indices of labor and capital inputs, and a measure of the intermediate inputs electricity and fuels. Labor is the number of man-years hired for production, and firms distinguish between their blue- and white-collar workers. Gross output, capital, materials, electricity, and fuels each have their own price deflator and are each measured in real 1980 Chilean pesos. We provide more details on the variables in Appendix B.

Table 1 provides some macroeconomic background as well as some summary statistics for the industries we examine. The first year of the data is 1979. By 1979, most of Pinochet's economic policies were already in place. The Latin American debt crisis led to a recession in 1982-83 during which industrial output and employment fell. Industrial output rose again in 1984, stalled in 1985, and then continued to rise throughout the decade. These macroeconomic cycles are apparent in the first column of Table 1 where real GDP is reported for 1979-86. When we address estimation issues in the next section, we will take these macroeconomic cycles into account by allowing the  $\omega_t(l_t, k_t)$  function to be different for each of these different time periods (i.e.  $t=1,2,3$ .)

It is also evident from Table 1 that this period is characterized by major consolidation and shake out; the number of plants falls in every industry from the beginning to the end of the sample. There is, though, both entry and exit in our sample. Although the original work by Olley and Pakes devoted significant effort to highlighting the importance of not using an artificially balanced sample (and the selection issues that arise with the balanced sample), we simply note that our sample is not balanced, and we do not focus on selection issues.<sup>18</sup>

Table 2 reports the percentage of firm-level observations reporting non-zero levels of investment, fuels, materials, and electricity. In industries with adjustment costs, many firms may not invest every period. However, the methods of Olley and Pakes revolve quite centrally around observing non-zero investment decisions because no proxy is available for firm/year observations with zero investment. Instead, these firms are truncated from the estimation routine.<sup>19</sup> They use data from the U.S. Census of Manufacturers, and they find that 8% of firm/year observations are reported to be zero; accordingly, they truncate these observations.

In our data, the investment problem is much more severe; in seven of the eight industries, we would have to truncate over 50% of the firm-level observations. In contrast, almost no firm in all

---

<sup>18</sup> Technically speaking, there is still a selection issue even with unbalanced data, but it is unlikely to be very important from an empirical viewpoint. We have experimented with dual index selection corrections of Olley and Pakes and found that they made little difference. In order to focus on the intermediate inputs issue, we do not include those methods or results in this paper.

<sup>19</sup> See Olley and Pakes (1996) for a detailed explanation of this problem. Briefly, it arises because productivity is not a smooth function of investment at zero investment levels.

eight industries reports zero for materials or electricity usage; if we can justify using either one of these intermediate inputs, we will not need to truncate observations.

We choose electricity as our primary proxy for productivity, although we also experiment with using materials and fuels. One reason for choosing electricity is that it is an input that all firms need; we observe positive use of electricity in every year for almost every firm in the Census. Additionally, almost no firm in all eight industries reports that it generates electricity, or that it sells electricity, and we interpret this observation (together with some common sense) as an inability of firms to store electricity. The inability to store electricity means that its use should be highly correlated with the year-to-year productivity terms and hence be a good proxy. Materials and fuels, on the other hand, are probably easy to store over time, and hence new purchases of these inputs (which is what we observe in the data) may not exactly track the amount of these inputs used in the production process in a particular year.

Another reason for choosing electricity (relative to materials) relates to using *Condition 3*; if one chooses to invoke the separability assumption in the second stage, we feel it seems less restrictive when it is imposed on an input that accounts for about one percent of output value as opposed to an input that accounts for about half of the value of output. (The revenue shares of the various inputs are given in Table 3.) Alternatively, it seems more intuitive that electricity (or fuels) might actually be separable whereas materials could represent a lot of possible inputs— some of which may interact with other inputs in a manner that violates separability. We next turn to estimation issues, where we explicitly treat electricity as the proxy, although we later experiment with other inputs.

## 5. Estimation

Estimation proceeds in two stages, each of which is comprised of several steps. In the first stage, we use a partially linear model to estimate the coefficients on the freely variable factors: skilled and unskilled labor, materials, and fuels. Electricity is our primary proxy variable and hence is not included in this list; the estimate for electricity will come from the second stage.

We develop two approaches to the second stage. The first approach, which we call the unrestricted approach, combines the consistent estimates of the coefficients on the variable factors from the first stage with two assumptions: i) the non-forecastable component (or innovation in) productivity is uncorrelated with capital; and ii) the productivity process is first order Markov. The second approach, which we call the restricted approach, can be used when one observes revenue shares and one believes the production technology is separable in the proxy input.

The unrestricted case has the obvious appeal that it does not impose a constraint that may be false in the data. Were we to estimate just a few production functions, the computational costs associated with the 11 step procedure that nests a two dimensional grid search (because the objective function is not globally concave) would not be significant. However, in the results section of this paper we need to estimate production functions for each of eight different industries using eight different estimators. In addition, we are going to bootstrap our standard errors, drawing from the empirical distribution function 30 times, each time estimating these same 64 production technologies. Hence we wish to run our estimation routine almost 2000 times. The computational cost of doing so with the unrestricted approach is currently too high to undertake. The restricted approach has the advantage of significantly reducing the computational costs associated with estimation. Our first set of results then explores the plausibility of this restriction.

In this section, we cover only the specifics and intuition of how our estimation routine is implemented and why it “works.” Consistency proofs for our estimators would use results from Pakes and Olley (1995). Since we bootstrap our standard errors, we do not need to impose the additional assumptions necessary to use asymptotics for the standard errors. We begin with a discussion of each stage of the estimation routine in turn. We then describe the bootstrap approach that we use in order to estimate the standard errors. Readers interested in implementing our estimation routine are also directed toward the Estimation Recipe in Appendix C. That appendix provides a detailed step-by-step guide to the entire estimation routine.

### *The First Stage*

To keep the exposition straightforward, we approximate the production technology

$$Y_t = f(K_t, L_t^s, L_t^u, M_t, F_t, E_t, \exp(\omega_t), \exp(\eta_t))$$

with a first order Taylor series (i.e. Cobb-Douglas,) or

$$y_t = \beta_0 + \beta_k k_t + \beta_s l_t^s + \beta_u l_t^u + \beta_m m_t + \beta_f f_t + \beta_e e_t + \omega_t + \eta_t, \quad (16)$$

where  $y_t$  is the log of gross output in year  $t$ ,  $k_t$  is the log of the plant’s capital stock,  $l_t^s$  is the log of skilled labor input,  $l_t^u$  is the log of the unskilled labor input, and  $m_t$ ,  $f_t$  and  $e_t$  denote log-levels of materials, fuels and electricity. (As discussed earlier, our approach generalizes to more flexible functional forms.)

We rewrite (16) as

$$y_t = \beta_s l_t^s + \beta_u l_t^u + \beta_m m_t + \beta_f f_t + \phi(e_t, k_t) + \eta_t, \quad (17)$$

where

$$\phi_t(e_t, k_t) = \beta_0 + \beta_k k_t + \omega_t(e_t, k_t).$$

Equation (17) is partially linear; it is linear in variable inputs, and non-linear in electricity and capital. The goal in this first stage is to obtain estimates on the coefficients of the inputs that enter (17) linearly (i.e.  $\beta_s$ ,  $\beta_u$ ,  $\beta_m$ , and  $\beta_f$ .) Here we follow the general approach for semiparametric estimation outlined in Robinson (1988).<sup>20</sup>

This approach requires us to construct estimates of the conditional moments  $E(y_t|k_t, e_t)$ ,  $E(l_t^u|k_t, e_t)$ ,  $E(l_t^s|k_t, e_t)$ ,  $E(m_t|k_t, e_t)$ , and  $E(f_t|k_t, e_t)$ . We do this by projecting  $y_t$  (for example) on  $k_t$  and  $e_t$  using a locally weighted quadratic least squares approximation. Readers not familiar with local regression smoothing might find it helpful to think of this step as using weighted least squares to construct predictions of  $y_t$  given  $(k_t, e_t)$  using a second order approximation in  $(k_t, e_t)$  (a quadratic in  $(k_t, e_t)$ ). For any particular point  $(k_t^*, e_t^*)$  for which an estimate of the expected value of  $y_t$  is necessary, the regression weights the observations closest to the point  $(k_t^*, e_t^*)$  most heavily. A consistent estimator of  $E(y_t|k_t = k_t^*, e_t = e_t^*)$  is then the intercept from this local quadratic regression.

We can write the expectation of (17) conditional on  $(k_t, e_t)$  as:

$$E(y_t|k_t, e_t) = \beta_0 + \beta_s E(l_t^s|k_t, e_t) + \beta_u E(l_t^u|k_t, e_t) + \beta_m E(m_t|k_t, e_t) + \beta_f E(f_t|k_t, e_t) + \phi(e_t, k_t) \quad (18)$$

where we have made use of the fact that  $E(\eta_t|k_t, e_t) = 0$  and  $E(\phi(e_t, k_t)|e_t, k_t)$  is itself. We subtract (18) from (17) to obtain:

$$\begin{aligned} y_t - E(y_t|k_t, e_t) &= \beta_s (l_t^s - E(l_t^s|k_t, e_t)) + \beta_u (l_t^u - E(l_t^u|k_t, e_t)) \\ &\quad + \beta_m (m_t - E(m_t|k_t, e_t)) + \beta_f (f_t - E(f_t|k_t, e_t)) + \eta_t \end{aligned} \quad (19)$$

This differences  $\phi_t(\cdot)$  out of the estimating equation. And since  $\eta_t$  is, by assumption, conditionally mean independent of the variable inputs, no-intercept Ordinary Least Squares can be used to obtain estimates of the parameters on all of the variable coefficients. Note that both the dependent and

---

<sup>20</sup> See section 5.2, and especially the first few pages of 5.2.1 of Pagan and Ullah (1999), for a (relatively) understandable and (very) helpful discussion of kernel-based estimates of the coefficients on the linear terms of the estimating equation.

explanatory variables are constructed variables that depend upon the local least squares estimates, making the OLS standard errors incorrect for our application. Our standard errors will account for all of these sources of variance, but we defer discussion of their computation until later in this section.

This completes the first stage of the estimation routine. Although there are several steps, no single step is more complicated than a locally weighted least squares regression.<sup>21</sup> If we were only concerned with the marginal productivities of the variable inputs (except the coefficient on the proxy variable), we could stop here. In general, one may want to recover a plant-level measure of productivity or an estimate of returns to scale, and for this exercise we need a second stage to which we next turn.

### *The Second Stage (Not Assuming Separability)*

The estimation routine requires a second stage in order to identify an estimate of  $\beta_e$  and  $\beta_k$  because electricity and capital enter  $\phi(\cdot)$  twice, i.e.  $\phi_t(e_t, k_t) = \beta_0 + \beta_e e_t + \beta_k k_t + \omega_t(e_t, k_t)$ , and hence are not identified without further restrictions. In this section we outline the unrestricted approach. In the next section, we briefly describe how one could proceed if separability in the proxy input holds (the restricted approach.)

Identification of these coefficients obtains in part from assuming that capital, a state variable, is slow to adjust to the productivity shock. While capital might adjust to the expected part of a productivity shock, the identifying assumption maintains that capital does not instantaneously adjust to the unexpected part of the productivity shock. To operationalize this notion, we must impose some structure on the stochastic process of the transmitted productivity shock,  $\omega$ . We assume that  $\omega_t$  follows a first order Markov process.<sup>22</sup> In particular, we can write

$$\omega_t = E(\omega_t | \omega_{t-1}) + \xi_t, \tag{20}$$

where  $\xi_t$  is the “news” in the transmitted shock.

The two moment conditions that identify the coefficients are then given by:

$$E(\xi_t + \eta_t | k_t) = E(\xi_t | k_t) + E(\eta_t | k_t) = 0, \tag{21}$$

---

<sup>21</sup> An alternative approach to this stage is to use a polynomial approximation for  $\phi(e_t, k_t)$ . We have done this (allowing for different sub-periods of the sample according to macroeconomic cycles) and we find that a third order polynomial approximation gives very similar estimates of the parameter estimates and of  $\phi(e_t, k_t)$ .

<sup>22</sup> The estimation routine easily generalizes to allow for a higher order Markov process.

and

$$E(\xi_t + \eta_t | e_{t-1}) = E(\xi_t | e_{t-1}) + E(\eta_t | e_{t-1}) = 0. \quad (22)$$

The first moment condition states that capital does not respond to the innovation in productivity. The second moment condition states that last period's electricity choice is uncorrelated with the innovation in productivity this period. We employ a Generalized Method of Moments estimator to find the parameter estimates that most closely match our samples to these population moment analogs. The steps we take to do so now follow.

The estimation algorithm begins by choosing a starting value for  $(\beta_e, \beta_k)$ . We denote these candidate values by  $(\beta_e^*, \beta_k^*)$ . For any candidate values, we can re-write (16) to yield:

$$y_t - \beta_s l_t^s - \beta_u l_t^u - \beta_m m_t - \beta_f f_t - \beta_e^* e_t - \beta_k^* k_t = \omega_t + \eta_t \quad (23).$$

Substitution using (20) yields:

$$y_t - \beta_s l_t^s - \beta_u l_t^u - \beta_m m_t - \beta_f f_t - \beta_e^* e_t - \beta_k^* k_t - E(\omega_t | \omega_{t-1}) = \xi_t + \eta_t \quad (24)$$

Everything on the left-hand-side of (24) is either data or was estimated in the first stage except for the term  $E(\omega_t | \omega_{t-1})$ . Put another way, if we “knew” the conditional expectation  $E(\omega_t | \omega_{t-1})$ , we could compute the residual that enters the moment condition (22), i.e.  $\xi_t + \eta_t$ .

The next step is to obtain an estimate of  $E(\omega_t | \omega_{t-1})$ . We begin by noting that,

$$E(\omega_t | \omega_{t-1}) = E(\omega_t + \eta_t | \omega_{t-1}).$$

Conditional on our candidate values  $(\beta_e^*, \beta_k^*)$ , (23) implies that:

$$\omega_t + \eta_t = y_t - \hat{\beta}_s l_t^s - \hat{\beta}_u l_t^u - \hat{\beta}_m m_t - \hat{\beta}_f f_t - \beta_e^* e_t - \beta_k^* k_t.$$

This gives us an estimate of  $\omega_t + \eta_t$ . From the first stage of estimation, we have

$$\omega_{t-1} + \eta_{t-1} = \phi_{t-1} - \beta_e^* e_{t-1} - \beta_k^* k_{t-1}$$

where  $\phi_{t-1}$  is estimated for each of the three time periods 1979-81, 1982-83, and 1984-86, effectively allowing the  $\omega_t(e_t, k_t)$  function to differ in arbitrary ways across the periods. We then use local quadratic least squares with our dependent variable as  $\omega_t + \eta_t$  and our explanatory variable  $\omega_{t-1} + \eta_{t-1}$  to estimate  $E_t(\omega_t | \omega_{t-1})$ .

With an estimate of  $E(\omega_t|\omega_{t-1})$  computed, we can now compute an estimate of the residual  $\xi_t + \eta_t$  using (24). We then use a grid search to minimize the GMM criterion function

$$\min_{\beta} \left[ \left( \sum_i \sum_{t=T_{i0}}^{T_{i1}} (\xi_{i,t} + \hat{\eta}_{i,t}) k_t \right)^2 + \left( \sum_i \sum_{t=T_{i0}}^{T_{i1}} (\xi_{i,t} + \hat{\eta}_{i,t}) e_{t-1} \right)^2 \right], \quad (25)$$

where  $T_{i0}$  and  $T_{i1}$  index the second and last period a firm is observed.<sup>23</sup>

It is perhaps helpful to note in a slightly different manner what the moment condition represents. The expectation of output less the contribution of inputs equals the error, or the transmitted shock plus another additive independent error,  $\omega_t + \eta_t$ . This error cannot be used in a moment condition because neither  $k_t$  nor  $e_{t-1}$  (nor  $e_t$ ) is mean independent of the error term. However, once we condition out the expected part of this error term (conditional on last period's productivity shock), the new error term  $\xi_t + \eta_t$  is mean independent of  $k_t$  and  $e_{t-1}$ . Our estimator exploits this condition. We now turn to the case when separability is assumed.

#### *The Second Stage (Assuming Separability)*

The procedure when separability is assumed is very similar to that outlined above. The two primary changes are: i)  $s_e e_t = \beta_e e_t$  is net out of  $y_t$  whenever the other variable inputs are net out, and ii) minimization is done just with respect to the capital coefficient. Also, only the moment condition relating capital is used (although in principal you could use both moments above to improve the efficiency of the estimator, or to test the overidentifying condition.) In short, conditional on a candidate value  $\beta_k^*$ , (23) now implies that:

$$\omega_t + \eta_t = y_t - \hat{\beta}_s l_t^s - \hat{\beta}_u l_t^u - \hat{\beta}_m m_t - \hat{\beta}_f f_t - s_e e_t - \beta_k^* k_t.$$

With an estimate of  $E(\omega_t|\omega_{t-1})$  computed as described above, we can again compute an estimate of the residual,  $\xi_t + \eta_t$ . We then use a grid search to minimize the GMM criterion function

$$\min_{\beta} \left( \sum_i \sum_{t=T_{i0}}^{T_{i1}} (\xi_{i,t} + \hat{\eta}_{i,t}) k_t \right)^2, \quad (26)$$

where  $T_{i0}$  and  $T_{i1}$  index the second and last period a firm is observed. We now turn to the standard errors.

---

<sup>23</sup> We experimented with the optimal weighting matrix and found it had little effect on either our estimates or the standard errors.



### *Computation of Standard Errors*

While the previous sections explain how the parameters are estimated using intermediate inputs as a proxy, we did not discuss the estimation of the standard errors. Estimation of these standard errors requires us to account for every source of variance in every estimator that enters our routine. We now summarize them. In the first stage, five preliminary local quadratic least squares estimates are used in a no-intercept OLS estimator to obtain estimates of coefficients on the freely variable inputs. In the second stage, these estimated coefficients feed into each of four preliminary local quadratic least squares estimators, and both the estimated first-stage coefficients and estimates from these four local least squares estimators are combined to yield the final estimating equation for the capital coefficient. This coefficient is then obtained by using a Method of Moments estimator. In total, there are 11 estimating equations, and all of the preliminary estimators get used somewhere (not always just once) in the process of obtaining the final set of production function estimates.

These many intermediate steps introduce “noise” into the estimation routine. Additionally, each bit of noise is generated by the exact same sample; this makes specification of the covariances in the errors across equations a difficult task, but one that must be undertaken if one wants an estimate of the asymptotic standard errors. Pakes and Olley (1995) provide the theoretical details of how one would compute the asymptotic standard errors for such an estimation routine as ours; it essentially has a correction for the potential noise from every source of variance that feeds into the last estimating equation (equation (26).)

Instead of undertaking the difficult and somewhat arbitrary task of specifying and estimating variances and covariances of errors across equations, we choose to use the bootstrap, a simpler approach that is made feasible by the low cost of computing power. The bootstrap has an intuitive appeal to it; use the observed data as an approximation to the true underlying distribution of data, sample from it repeatedly, and compare the variability of estimates across these different samples. In cases where the data are i.i.d., sampling with replacement can be used to generate a “new” sample drawn from the actual observed sample. A collection of new samples can then be used to observe the variability in the estimator across samples, i.e. variance in the estimates due to sampling variability.

The standard deviation from this sample of estimates is the sampling variability, and there is no need to appeal to asymptotic normality; as long as the observed sample is a close approximation to the true underlying distribution function, all of the theory works. Sampling from the observed distribution of the  $(Output_{it}, Input_{it})$  vector has the advantage of allowing for arbitrary covariance between the errors and the regressors when the data are not i.i.d., as is certainly true with the

unbalanced plant-level data we see. However, when the data are not i.i.d. one must alter the sampling procedure in a way that allows the generated samples to reflect the likely true underlying distribution.

We now summarize our sampling approach, providing the details in the recipe section. We tailor our approach in a manner designed to accommodate the non-i.i.d. nature of our data. We employ a technique known as the block bootstrap, treating the entire “block” of observations on a firm (i.e.  $(Output_i, Input_i)$  across  $t$ ) as one observation. We choose a sampling procedure (described in the recipe in Appendix C) that has two parameters, and we calibrate these parameters until we closely match the 32 conditional means of annual number of firms and annual total value of output in the industry. After these parameters are set, we draw 30 samples for each industry, and we compute the OLS, fixed effects, instrumental variables, Olley-Pakes, and LP estimates for each sample. The variance covariance matrix of the estimated parameters is an estimate of the variance of the estimation routine. We also construct the differences between our estimator and each of OLS, fixed effects, and Olley-Pakes estimates across the bootstrapped samples; this gives us an estimate of the bias when the estimation routine is nested in our routine, as are the OLS, fixed effects, and instrumental variable estimators.

## 6. Results

In this section, we present several sets of results with several objectives in mind. Our goal is not just to show that one obtains statistically significant estimates of the parameters of the production function (although this is nice.) Rather, we wish to highlight how the estimators using intermediate inputs to control for unobservables differ in predictable and informative ways from other existing and commonly used estimators. We also show that these differences are remarkably robust across the eight industries we examine.

We begin by comparing production function estimates across the eight industries using the unrestricted and restricted approach with electricity as the proxy. Since the results are very similar across these two approaches, we then use the less computationally intensive approach (the restricted approach) for the rest of our results. We then compare our estimator to OLS, fixed effects, instrumental variables, and Olley-Pakes. Finally, while the bulk of our results are obtained using electricity as the as the proxy for the transmitted productivity shock, our last table compares results across proxies.

Table 4 presents parameter estimates of the restricted and unrestricted cases. In the unrestricted case, we estimate the coefficient on electricity. In the restricted case, this parameter is set equal

to electricity’s industry-level average revenue share.<sup>24</sup> The coefficients on the variable factors (excepting the proxy) are presented in the first four rows. These parameter estimates are obtained in the first stage of the estimation procedure. Because the first stage is the same for both the restricted and unrestricted cases, the parameter estimates (and standard errors) are identical. We have ordered the industries, represented by columns, in order of decreasing observations. That is, ISIC 311 has the most observations while 313 has the fewest. We find that almost all the coefficients are precisely estimated at conventional levels with the exception of the coefficient on fuels for some of the smaller industries.

Because we want to estimate almost 2000 production functions, we are especially interested in whether imposing the restricted approach (the cost-minimization condition) for electricity significantly changes the estimated coefficients on capital and electricity. We first consider the coefficient on capital. When our null hypothesis is that the capital coefficient from the restricted case equals that from the unrestricted case, we accept this null in seven of the eight industries. Only in ISIC 311 are the capital coefficients significantly different. For the case of the coefficient on electricity, in all eight industries the coefficient from the restricted case lies within two standard deviations of the revenue share. (The revenue share for electricity is listed in the row labeled “restricted.”) We conclude that we are not doing violence to the data by simply imposing that the coefficient on electricity equals its revenue share. As noted above, this greatly simplifies computation of the estimates. For the remainder of the results, we focus only on the restricted case.

We are especially interested in investigating whether our results are consistent with the story of a proxy controlling for a unobserved transmitted productivity shock. We begin this investigation by comparing our estimates to Ordinary Least Squares (OLS) estimates on the same data. Marschak and Andrews were concerned that the transmitted productivity shock would be positively correlated with variable inputs. In this case, OLS estimates of the coefficients on the variable inputs are likely to be biased upward. To the extent that capital also responds to the transmitted productivity shock, its estimated coefficient would also be upwardly biased. However, if capital is not correlated with this period’s transmitted shock (but variable inputs are), or capital is much less weakly correlated with the productivity shock than the variable inputs are, it is straightforward to show that the OLS estimate on capital will be biased downward (see Section 2.)

Figure 1 provides six histograms pertaining to industry ISIC 311. We initially focus our attention on the first histogram (top left corner.) This graph gives the empirical distribution of the difference

---

<sup>24</sup> We experimented with the median and an approach suggested in Griliches and Ringstad (1971) and found that (for the subset of results on which we compared) the different approaches yielded very similar results.

between the OLS estimate and our (LP) restricted estimate of the coefficient on unskilled labor,  $\beta_u$ . This histogram is constructed in the following manner. The production function is estimated 31 times (once using the full sample and then with 30 blocked sub-samples.) With each of these 31 samples, we also estimate the OLS coefficients. To generate the histogram, we simply count (across the samples) the number of times the difference between the OLS and LP estimate falls within a given range. If there were no difference between the OLS and LP estimates, we would expect to see the histograms for the variable factors centered symmetrically around zero; approximately 50% of the samples should yield differences to the right of zero. The histogram for unskilled labor shows that for the sample and for every subsample used in the block bootstrap, the OLS estimate exceeded the LP estimate. (All counts are to the right of zero.) The histograms for the other variable inputs show similar patterns. Indeed, in this industry, the OLS estimate is never less than the LP estimate for any variable factor for any of the subsamples on which the production function was estimated.

At the bottom left corner of Figure 1, we present the histogram of the differences between the OLS and LP estimates for the capital coefficient. This histogram shows that for all 31 times we estimated the production function, the OLS estimate of  $\beta_k$  was smaller than the LP estimate. Taken together with the histograms for the variable factors, the evidence substantiates Marschak and Andrews' concern and suggests that our input proxy is doing what it is intended to do— control for an unobservable transmitted shock that appears to be highly correlated with freely variable inputs.

The histogram in the lower right corner gives the empirical distribution of the difference in estimated returns to scale between the OLS and LP estimates. Because OLS overestimates the coefficients on the variable factors and underestimates the coefficient on capital, it is, in general, not obvious whether OLS will over- or under-estimate returns to scale. It's conceivable the over-estimated variable input coefficients and underestimated capital coefficient would just cancel each other out. In ISIC 311, this is not the case. Rather, OLS consistently overestimates returns to scale.

The pattern of biases in Figure 1 is the same as that found by Olley and Pakes in the single industry they examined (and using investment as their proxy.) We wish to examine whether this pattern holds consistently across all eight industries. Presenting 48 histograms is unwieldy. Instead, Table 5 summarizes the histograms for all coefficients (and returns to scale) for all eight industries. The first row for each industry gives the mean difference ( $\hat{\beta}_{OLS} - \hat{\beta}_{LP}$ ) between the OLS and LP

estimates of each coefficient in the production function as well as the difference in estimated returns to scale.

These industries are again ordered by the number of observations. ISIC 311 has the most observations and ISIC 313 has the fewest. In all industries, we draw 30 blocked sub-samples on which we then estimate both OLS and LP estimates. For example, in ISIC 381, the mean difference between the OLS and LP estimates is positive for all the variable factors. For the case of skilled labor, only 58 percent of our samples yielded OLS estimates that were greater than the LP estimates while for unskilled labor, materials, and fuels, the OLS estimates were greater in 100, 100, and 84 percent (respectively) of the sub-samples over which we estimated the production function. The coefficient on capital was smaller with LP than with OLS in 97 percent of the samples (30 of the 31.) While none of the industries yields results as stark as those for ISIC 311, several general patterns emerge in the OLS versus LP comparison. In general, we find that OLS usually overestimates the coefficients on variable factors, and OLS always underestimates the coefficient on capital. The magnitude and statistical precision of these difference varies by factor and by industry, but the flavor of the results tend to be most convincing in the industries with the most observations (which have the smallest sampling error.)

Ordinary Least Squares is but one alternative to our approach. Other alternatives that are used include Fixed Effects, Instrumental Variables, and the Olley Pakes estimators. We compare our estimates to those obtained with all of these alternatives. In order to keep the quantity of results manageable, we present a handful of summary comparisons.<sup>25</sup> The first of these comparisons is given in Table 6.

In Table 6, we test the null hypothesis that the results from the LP estimator and an alternative estimator are (overall) consistent with one another. The test asks whether any of the five estimated parameters is significantly different across the two approaches, and thus has a similar flavor to that of the standard F-test in a regression framework, where the null is that all of the coefficients are zero (here it is differences) against the alternative that at least one is different from zero. Hence, if the difference between the LP estimate and the alternative estimate of one or more coefficients in the production function is statistically significantly different from zero, we reject the null hypothesis. The cell entries in Table 6 give the percent of industries for which we reject this null hypothesis. To fix ideas, consider the first column of Table 6. At the 5% significance level, we reject this null hypothesis in 6 of the 8 (or 75%) industries. At the 10 and 20 percent significance levels, we reject

---

<sup>25</sup> Complete results for all the alternative estimators are available on request.

the null hypothesis in all eight industries. (It is important to keep in mind that the standard errors of the estimates account for the “noise” introduced by every step of the estimation routine and for the fact that rates of convergence for nonparametric regressions are typically slower than that of their parametric ones. In the testing sense, “significant” results are thus harder to obtain when less structure is placed on the problem.)

Just as we compared the LP estimates to the OLS estimates, we also compare the LP estimates to Fixed Effects estimates. These results, for the null hypothesis that all coefficients are the same with fixed effects as with LP, are given in the second column of Table 6. There we note that at all three reported significance levels, we reject the null hypothesis in all of the industries. One way to interpret the results in the first two columns of Table 6 is to think of OLS and Fixed Effects as nested restrictions on the stochastic process of productivity. In particular, OLS imposes that  $\omega_{it}$  be i.i.d. whereas the fixed effects estimator forces  $\omega_{it}$  to be constant across time and allows it to vary only across firms. Both the OLS and Fixed Effects assumptions are nested restrictions of our more general first order Markov assumption. We soundly reject both nested restrictions.

The third column considers an instrumental variables (IV) estimator as the alternative to LP. The “instruments” we use to construct this estimator are last period’s variable inputs. These are, at best, very questionable instruments given the strong serial correlation in input usage and in productivity. Nonetheless, these instruments are frequently employed, and thus we investigate whether these IV estimates are significantly different from the LP estimates. We reject that the difference between the IV estimates and the LP estimates is zero for half the industries at the 5 percent significance level and for 6 and 8 of the industries at the 10 and 20 percent significance levels respectively.

One cannot determine from Table 6 *how* the OLS, FE, and IV estimates differ from the LP estimates. Table 5 addresses this exact issue for the case of OLS. Tables analogous to Table 5 but for FE and IV estimates are available in the Appendix. The general pattern with OLS and IV is that displayed in Table 5 and is exactly that predicted by the theory; coefficients on the variable inputs are usually overestimated while the coefficient on capital is underestimated relative to the LP estimates. With fixed effects, there is no clear theoretical prediction, but the results are significantly (almost strikingly) different.

We also investigate whether the Olley Pakes (OP) estimates are the same as the LP estimates. Doing so is somewhat subtle. Recall, the OP approach can only be estimated on the set of firms reporting strictly positive investment. The last three columns of Table 6 address aspects of the OP versus LP comparison. Column 4 reports the OP versus LP comparison when we restrict the sample

to those firms that report strictly positive investment. Column 5 addresses the issue of selection bias when one truncates on firms that report positive investment. In that column, we compare the difference between the LP estimates when the entire sample is used and the LP estimates when we only use those firms that invested. If there were no selection bias and plants that reported positive investment were a random draw from the population of plants, then the LP estimates from the full sample and those from the positive investment subsample should be the same. In the last column, we compare the estimates from OP versus LP– estimates which are based on different samples. Each of these columns is discussed in turn.

Column 4 is a comparison of OP versus LP when we restrict the sample to positive investment plants. Even if there were truncation bias, it is not obvious whether such truncation bias would impact the OP versus LP comparison in any particular direction. At the 5 percent level, we reject that OP and LP give the same parameter estimates in 3 of the 8 industries. At the 10 and 20 percent levels, we reject the null for 5 and 7 of the 8 industries respectively. We return to a more in-depth discussion of how the OP and LP estimates in this column differ below.

Column 5 directly addresses the issue of truncation bias. There we ask whether we obtain the same parameter estimates when we truncate the sample based on only positive investment plants. In half the industries (at the 5 and 10 percent significance levels) we reject the null hypothesis of no difference in the estimated coefficients. At the 20 percent level, we reject the null in 7 of the 8 industries. We take this to suggest that there probably is some truncation bias in our sample, but the evidence is not overwhelming. The sign of the difference between the estimates with the full sample and the estimates with the truncated sample does not seem particularly salient to the issues we are investigating.

Column 6 simply compares the OP and LP estimates and asks whether they are the same. We reject the null in 4, 5 and 8 of the 8 industries at the 5, 10, and 20 percent significance levels. Clearly, OP and LP estimators give different results, but we are unable to identify why these estimates differ. Truncation bias matters (column 5), but even absent truncation bias (column 4), the OP and LP estimates are different.

Table 7 revisits all of the comparisons across different estimators but does the accounting differently. Instead of asking whether *any* coefficient in the production function is different with different estimators, here we simply count how many of the five estimated parameters are different across all eight industries. To fix ideas, consider the OLS versus OP comparison. With eight industries, there are 40 pairs of estimated parameters. The first column of Table 7 reports the percentage of these 40 pairs for which we can reject that the OLS and LP estimates are the same. In other words, at

the 5% significance level, the difference between the OLS and LP estimates is significantly different from zero in 32.5 percent of the 40 cases. The general message from the first three columns of Table 7 is that the LP estimates differ from the more traditional estimates (OLS, FE, and IV) in about 40 to 60 percent of the cases depending on the level of statistical significance adopted. This corresponds to 2 to 3 of the 5 coefficients in the production function differing when one uses LP instead of the more traditional estimates. The message is similar but less pronounced in the last three columns of Table 7 where aspects of the OP versus LP comparison are explored.

Table 8 addresses the comparison of estimated returns to scale across the various estimators. We find that OLS, Fixed Effects, and IV estimators give different estimated returns to scale than does the LP approach. Although it is not apparent in the table, all these alternatives to LP give inflated returns to scale relative to LP. The comparisons of OP versus LP suggest that both approaches give about the same estimated returns to scale. There is weak evidence that what differences do appear may be due to the truncation bias of selecting on positive investment plants.

Recall our motivation for introducing the intermediate input proxy in the first place. We were concerned that the transmitted component of the productivity shock may have a forecastable component and that if capital has already adjusted to this component of the shock, investment will only respond to the “news” in the shock. The freely variable inputs, on the other hand, will respond to both the news and the forecastable component. To the extent that this might be true in the data, the OP estimates of the coefficients on some variable factors might be biased upward and the OP estimates of the capital coefficient might be biased downward. We investigate the extent to which this is true in Table 9.

In Table 9, we examine the difference between the OP estimates and the LP estimates when we only use data from plants reporting positive investment. We continue to order the industries in decreasing order based on the number of observations. Especially for the industries that have more observations, we find that the differences between the OP estimates and the LP estimates are broadly consistent with the concern that motivated the use of the intermediate input proxy. Consider, for example, ISIC’s 311 and 381. In these industries, the mean difference between the OP estimate and the LP estimate on the variable factors is positive (i.e. the OP estimate is higher) for 7 of these 8 coefficients. The percentage of the bootstrapped samples for which this is true is large— usually between 70 and 100 percent. Overall, the mean difference between the OP and LP estimates (with the truncated sample) is non-negative for 24 of the 32 estimated coefficients on the variable factors. (The difference is zero for 3 coefficients.)



For the case of the coefficient on capital, if the OP and LP estimates of  $\beta_k$  were really about the same, we would expect the difference between them to be positive in 50 percent of the bootstrapped subsamples and negative in 50 percent. This is not the case. Rather, in 7 of the 8 industries, the difference is positive less than 50 percent of the time. That is, in 7 of the 8 industries, the OP estimate is lower than the LP estimate in more than half the sub-samples. However, the mean difference between the OP and LP estimate is positive in half the industries and negative in half. Thus the evidence appears mixed on the capital coefficient.

Summarizing, we soundly reject that the intermediate input proxy approach gives the same results as the more traditional OLS, IV, or Fixed Effects estimators. Further, the ways in which these traditional estimators and our estimator differ are predictable and consistent with the economics underlying our approach. Put another way, Marschak and Andrews were right. The estimates from our approach and that of Olley and Pakes also differ, but the comparison is not as dramatic as that between our estimator and the more traditional estimators. This is not surprising. The Olley and Pakes approach is itself a dramatic and creative improvement over the traditional approaches, while our approach is a more subtle change in implementation vis a vis Olley and Pakes. Nonetheless, the intermediate input proxy approach seems to work well. Furthermore, it is easy to implement, allows the researcher to use more of the existing data, and appears to address some situations in which the OP approach may not work well.

In our last table, Table 10, we explore the extent to which other inputs also proxy for the transmitted productivity shock. In all the results reported above, we used electricity as the proxy. In Table 10, we report the resulting coefficients on skilled and unskilled labor and on capital when we use either materials or fuels as the proxy. We report results for the largest four industries in an effort to keep the quantity of results manageable. Also, in order to be able to compare estimates across different proxies, we estimate the production functions only on the set of observations for which data are available for all the proxies. (This is the reason the results for electricity differ slightly from those in Table 4.) In all four industries, the coefficients on the variable factors are quite stable across the different proxies. In one of the four industries (ISIC 381), the coefficient on capital increases dramatically when materials is the proxy. In the other three industries, the coefficient on capital is relatively stable across proxies.

This paper has not explored any of the implications of the production function estimates. Rather, we have focused on documenting how our estimates of the production function parameters differ in substantive and predictable ways from estimates obtained with other estimators. More interestingly, perhaps, is that the biases associated with the traditional OLS approach appear to

carry over to most of the other estimators when the intermediate input proxy is not used. Thus it is very likely that different estimates of the coefficients in the production function will give rise to different estimates of plant-level productivity. In a separate paper, Levinsohn and Petrin (1999), we explore the implications of our estimation strategy for productivity dynamics. There we show that OLS consistently overestimates positive productivity gains and also consistently predicts larger falls in productivity when productivity is negative. To the extent that this bias is not corrected for in the other estimators, we should expect similar mistakes when computing productivity estimates. Alternatively, the results presented there suggest that properly accounting for the endogeneity of variable inputs may well matter when it comes to analyzing the implications of our estimates for productivity and (perhaps) for returns to scale.

## **7. Conclusions and Caveats**

[Forthcoming.]

## References

- Basu, S., and Fernald, J. (1995). Are apparent productive spillovers a figment of specification error. *Journal of Monetary Economics*, 36(1), 165–188.
- 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.
- Dunne, T., Roberts, M., and Samuelson, L. (1988). Patterns of firm entry and exit in u.s. manufacturing industries. *Rand Journal of Economics*, 19(4), 495–515.
- Griliches, Z., and Ringstad, V. (1971). *Economies of Scale and the Form of the Production Function*. North Holland.
- Griliches, Z., and Mairesse, J. (1995). Production functions: The search for identification. NBER Working Paper 5067.
- Heckman, J. (1981). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In Manski, and McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press.
- Klette, T. J., and Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11, 343–361.
- Levinsohn, J. (1998). Employment responses to international liberalization in Chile. Forthcoming in the *Journal of International Economics*.
- Levinsohn, J., and Petrin, A. (1999). When industries become more productive, do firms? investigating productivity dynamics. NBER Working Paper 6893.
- Lui, L. (1991). *Entry-Exit and Productivity Changes: An Empirical Analysis of Efficiency Frontiers*. Ph.D. thesis, University of Michigan.
- Lui, L. (1993). Entry, exit, and learning in the Chilean manufacturing sector. *Journal of Development Economics*, 42, 217–242.
- Lui, L., and Tybout, J. (1996). Productivity growth in Colombia and Chile: Panel-based evidence on the role of entry, exit and learning. In Roberts, M., and Tybout, J. (Eds.), *Producer Heterogeneity and Performance in the Semi-Industrialized Countries*, chap. 4. World Bank.
- Marschak, J., and Andrews, W. (1944). Random simultaneous equations and the theory of production. *Econometrica*, 12(3-4), 143–205.
- Olley, S., and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263–1298.
- Pagan, A., and Ullah, A. (1999). *Nonparametric Econometrics*. Themes in modern Econometrics. Cambridge University Press.
- Pakes, A. (1996). Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interaction. In Sims, C. (Ed.), *Advances in Econometrics, Sixth World Congress, Volume II*, pp. 171–259 New York. Cambridge.
- Pakes, A., and Olley, S. (1995). A limit theorem for a smooth class of semiparametric estimators. *Journal of Econometrics*, 65(1), 292–332.
- Pavcnik, N. (1999). Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. Working Paper, Department of Economics, Dartmouth College.
- Roberts, M., and Tybout, J. (1997). The decision to export in Colombia: An empirical model of entry with sunk costs. *American Economic Review*, 87(4), 545–64.
- Robinson, P. (1988). Root-n consistent semiparametric regression. *Econometrica*, 55(4), 931–954.
- Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics*, 39(3), 312–320.
- Tybout, J., de Melo, J., and Corbo, V. (1991). The effects of trade reforms on scale and technical efficiency: New evidence from Chile. *Journal of International Economics*, 31, 231–250.

## Appendix A

In this appendix we consider the use of intermediate inputs as proxies for productivity when firms operate in a competitive environment. We show the general conditions on the production technology which yield an intermediate input demand function  $\iota(\omega; p_L, p_\iota, K)$  that is strictly increasing in productivity ( $\omega$ ) (the price of output is normalized to 1.) This result permits the use of  $\omega(\iota, K)$  as an index for productivity.

*Definition.* An industry is competitive if firms take input prices and the output price for the homogeneous good as given.

Intermediate inputs are available as proxies in some imperfectly competitive environments, although the proof depends on the specifics of the competition. Proofs in an imperfectly competitive environment will likely rely on arguments from the literature on monotone methods.

*Assumption* The firm production technology  $Y = f(K, L, \iota, \omega) : R^4 \rightarrow R$  is twice continuously differentiable in labor ( $L$ ) and the intermediate input ( $\iota$ ), and  $f_{L\omega}$ ,  $f_{\iota\omega}$ , and  $f_{\iota L}$  exist for all values  $(K, L, \iota, \omega) \in R^4$ . The industry is competitive, and either a) this period's investment does not respond to this period's productivity, or b) it does not enter this period's capital. Productivity is observed before the choice of labor and the intermediate input are made.

The differentiability of  $f(\cdot)$  can be relaxed with the appropriate appeal to monotone methods. We treat capital as fixed, and assume both labor and the intermediate input respond to the productivity. With some additional complexity, it is possible to show the following result when capital also responds to  $\omega$ , and when more than one type of labor exists.

*Result.* Under the Assumption, if  $f_{\iota L}f_{L\omega} > f_{LL}f_{\iota\omega}$  everywhere, then  $\iota(\omega; p_L, p_\iota, K)$ , the intermediate input demand function, is strictly increasing in  $\omega$ .

*Proof.* Given the assumption, a profit-maximizing firm has an intermediate input demand function that satisfies

$$\text{sign}\left(\frac{d\iota}{d\omega}\right) = \text{sign}(f_{\iota L}f_{L\omega} - f_{LL}f_{\iota\omega})$$

(see Varian (1992), pp. 494-495.) Under mild regularity conditions on  $f(\cdot)$  that insure the Fundamental Theorem of Calculus holds for  $\iota(\cdot)$ ,

$$\iota(\omega_2; p_\iota, p_L, K) - \iota(\omega_1; p_\iota, p_L, K) = \int_{\omega_1}^{\omega_2} \frac{d\iota}{d\omega}(\omega; p_\iota, p_L, K)P(d\omega|K).$$

Since  $f_{\iota L}f_{L\omega} > f_{LL}f_{\iota\omega}$  everywhere, it follows that

$$\int_{\omega_1}^{\omega_2} \frac{d\iota}{d\omega}(\omega; p_\iota, p_L, K)P(d\omega|K) > \int_{\omega_1}^{\omega_2} 0 P(d\omega|K) = 0,$$

so

$$\iota(\omega_2; p_\iota, p_L, K) > \iota(\omega_1; p_\iota, p_L, K) \text{ if } \omega_2 > \omega_1.$$

## Appendix B

### *Background on Chile*

In 1973, Pinochet came to power and the economic environment changed dramatically. Fiscal austerity and a more outward-oriented set of economic policies was introduced. This liberalization included privatization of firms previously held by the government, relaxation of some 3000 government controlled prices, liberalized financial markets, a more market-driven labor market, removal of quantitative trade restrictions and a drastic reduction in tariff levels. These liberalizing measures were accompanied by contractionary macroeconomic policies which, when coupled with a decline in copper prices and the oil shocks, led to a severe recession that lasted through 1975. A recovery from 1976 to 1981 followed, and the broad picture is one in which liberalization was drastic and mostly complete by 1979, the first year of the data set. The Latin American debt crisis led to another recession in 1982-83 during which industrial output and employment fell. Industrial output rose again in 1984, stalled in 1985, and then continued to rise throughout the decade; see Table 1.<sup>26 27</sup>

### *Variable Construction*

Real value of output is computed by deflating the total annual sales revenues of a firm with an industry level price deflator constructed by the Banco Central de Chile. This deflator will control for changes in output prices over time arising from inflation. It will also control for changes occurring because of industry level demand shocks. Having controlled for these time-varying effects, we then rely on price-taking behavior at the firm-level to get comparable quantities across both firms and time.<sup>28</sup>

Construction of the real value of capital is documented in Lui (1991). Essentially, it is a weighted average of the peso value of depreciated buildings, machinery, and vehicles, each of which is assumed to have a depreciation rate of 5%, 10%, and 20% respectively. Alternatively, each type of  $K$  evolves according to

$$K_{jt} = (1 - \delta_j)K_{j,t-1} + i_{jt},$$

and the total capital index is then given by

$$K_t = \sum_j K_{jt}.$$

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

---

<sup>26</sup> In the mid-1980's, Chile's credit markets were such that smaller firms were alleged to have found credit difficult to obtain. While this is difficult to document in our data set, it implies that entry and exit patterns may have been driven by more than the relatively straightforward operate versus shutdown decision in our model. Hence, credit market issues suggest that firm size may have mattered in a way that our model does not accommodate.

<sup>27</sup> GDP data and related figures are available on-line in the Penn World Tables, Mark 5.6 at <http://www.nber.org>.

<sup>28</sup> See Griliches and Mairesse (1995) for a nice discussion of the failure of "one price." The industries we look at have hundreds of firms, and this observation provides us with some comfort that price-taking behavior, even as an approximation to reality, is a reasonable assumption. However, we do remain concerned about the potential for differences in output prices within the 3-digit industry level that might arise because of imperfect competition, especially that caused by product differentiation (e.g. differences in type and quality of output.) Because we do not observe firm-level output prices, we are not in a position to address this concern without placing further restrictions on the framework. See Griliches and Mairesse (1995) for a brief discussion of the econometric consequences of the failure of "one price." Klette and Griliches (1996) provide one set of restrictions that allows for a particular kind of correction for unobserved firm-level prices.

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.

## Appendix C

The purpose of this appendix is to provide a step-by-step guide on how to estimate production functions using intermediate inputs to control for unobservables. Here we document the approach taken when we use the assumption of separability to net out the contribution of the electricity proxy from output. The recipe below is not written with any particular software package in mind. In practice, we began with a combination of Stata and Gauss, but switched to S+ in the end because of the ease with which one can compute and view the results from the local least squares estimators.

### Estimation Recipe

#### Stage One:

1. Run locally weighted least squares regression of  $y_t$  on  $e_t$  and  $k_t$  to obtain an estimate of the function  $E(y_t|e_t, k_t)$ .
2. Run locally weighted least squares regression of  $l_t^s$  on  $e_t$  and  $k_t$  to obtain an estimate of the function  $E(L_t^s|e_t, k_t)$ .
3. Run locally weighted least squares regression of  $l_t^u$  on  $e_t$  and  $k_t$  to obtain an estimate of the function  $E(L_t^u|e_t, k_t)$ .
4. Run locally weighted least squares regression of  $m_t$  on  $e_t$  and  $k_t$  to obtain an estimate of the function  $E(m_t|e_t, k_t)$ .
5. Run locally weighted least squares regression of  $f_t$  on  $e_t$  and  $k_t$  to obtain an estimate of the function  $E(f_t|e_t, k_t)$ .
6. Construct  $Y(e_t, k_t) = (y_t - E(y_t|e_t, k_t))$  using the estimate of the conditional expectation from the local weighted least squares from step 1. This is the dependent variable in step 7. Similarly, difference out the predicted mean for each of the explanatory variables, and call these new regressors that are net of electricity and capital variation  $(X_1(e_t, k_t), X_2(e_t, k_t), X_3(e_t, k_t), X_4(e_t, k_t))$ .
7. Run no-intercept OLS regressing the constructed dependent variable  $Y$  on the vector of constructed independent variables  $(X_1, X_2, X_3, X_4)$ .

This completes the first stage of the estimation routine. Key estimated parameters from this stage are the production function parameters on all the variable inputs—  $\hat{\beta}_s, \hat{\beta}_u, \hat{\beta}_f,$  and  $\hat{\beta}_m$ .

#### Stage Two:

Here we discuss the unrestricted restricted estimator; the restricted estimator proceeds in a similar fashion, but without estimating  $\beta_e$ ; instead the average revenue share is used as an estimate of this parameter, and electricity's contribution  $s_t e_t$  is net out in step 1 and step 3 along with the other variable inputs.

1. Compute estimate of  $\phi_t(e_t, k_t)$  for each of the three different time periods 1979-81, 1982-3, and 1984-6 using local regression to predict  $y_t - \hat{\beta}_s l_t^s - \hat{\beta}_u l_t^u - \hat{\beta}_m m_t - \hat{\beta}_f f_t$  using  $(e_t, k_t)$  as explanatory variables. Save the estimate,  $\hat{\phi}_t(\cdot)$ .
2. Choose a candidate value for  $(\beta_e, \beta_k)$ , say  $(\beta_e^*, \beta_k^*)$ . A good starting value might be the OLS value from a Cobb-Douglas production function. (We use the robust but computationally expensive grid search, so a "good" starting value is not critical for us.)

3. Compute  $\omega_t \hat{+} \eta_t = y_t - \hat{\beta}_s l_t^s - \hat{\beta}_u l_t^u - \hat{\beta}_m m_t - \hat{\beta}_f f_t - \beta_e^* e_t - \beta_k^* k_t$ . For notation's sake, call the variable just computed "A".

4. Compute  $\omega_{t-1} \hat{+} \eta_{t-1} = \phi_{t-1} - \beta_k^* k_{t-1} - \beta_e e_{t-1}$ . Call this variable "B".

5. Regress "A" on "B" using locally weighted least squares. This generates a new variable. This new variable is constructed from the constant from the regression as it is evaluated at each observation. Call this new variable "C". "C" is  $E(\omega_t | \omega_{t-1})$ .

6. Using (24), and substituting "C" in for  $E(\omega_t | \omega_{t-1})$ , compute  $(\xi_t \hat{+} \eta_t)$ . This is the residual that is used in the moment equation. This residual is mean zero conditional on  $k_t$  and  $e_{t-1}$  at the "truth," and thus can be used to construct the sample analogs to the population moment conditions.

7. Using your favorite minimization routine, choose  $(\hat{\beta}_e, \hat{\beta}_k)$  to minimize the objective function from (26) (i.e. distance between the observed moments and zero.) This will entail iterations over the previous six steps.

### Estimation of Standard Errors

Our approach is based on the block bootstrap, where all firm-specific observations are considered en bloc (as one i.i.d. observation.) We call the two parameters of our bootstrap method  $\alpha_1$  and  $\alpha_2$ .

1. Sample with replacement from the blocks of firms that existed in either 1980 or 1981. Draw  $\alpha_1^*$  (number of firms in existence in 1980) times from this distribution. If a firm is selected, the entire block of observations related to that firm is included in the bootstrap sample.

2. Sample with replacement from the blocks of firm-level observations for those firms that entered in 1981. Sample  $\alpha_2^*$  (number of entrants in 1981) firms.

3. Repeat step two for each year 1982-1985, and pool all the observations from steps 1-3 into one sample. This is the bootstrapped sample.

4. Estimate the parameters using this bootstrap sample for OLS, fixed effects, Olley-Pakes, and our estimator.

5. Repeat 1-4 30 times, and use the standard deviation of the estimates as the estimate of the variability of the estimator.



TABLE 1  
Some Descriptive Statistics on Chilean Manufacturing

		Metals		Textiles		Food Products		Beverages	
Year	GDP	Plants	Value Added	Plants	Value Added	Plants	Value Added	Plants	Value Added
1979	997.6	459	10.0	503	12.4	1,537	39.0	211	14.1
1980	1,075.3	447	11.0	445	12.9	1,439	43.4	188	14.7
1981	1,134.7	413	11.5	403	11.3	1,351	42.7	158	10.4
1982	974.9	365	8.1	350	8.7	1,319	47.0	151	11.7
1983	968.0	322	8.3	327	9.7	1,297	42.9	148	13.1
1984	1,029.4	358	11.4	336	10.4	1,340	46.8	138	13.2
1985	1,054.6	351	9.6	337	10.8	1,338	49.1	127	11.5
1986	1,114.3	347	9.6	331	12.9	1,288	61.4	111	12.8

		Other Chemicals		Printing & Publishing		Wood Products		Apparel	
Year	GDP	Plants	Value Added	Plants	Value Added	Plants	Value Added	Plants	Value Added
1979	997.6	171	15.1	242	11.4	524	10.4	442	6.6
1980	1,075.3	166	16.9	227	10.5	449	8.7	398	6.7
1981	1,134.7	159	17.5	206	12.8	406	6.8	346	6.6
1982	974.9	148	14.8	196	8.9	358	6.5	305	5.3
1983	968.0	145	12.6	177	5.8	335	8.1	265	4.1
1984	1,029.4	151	12.9	167	5.8	339	10.3	294	6.4
1985	1,054.6	149	11.2	164	4.7	342	10.1	275	8.5
1986	1,114.3	153	9.1	163	4.5	313	5.3	280	11.3

Notes: GDP figures from the International Financial Statistics Yearbook. GDP and value added in millions of 1980 pesos.

TABLE 2  
Percent of Usable Observations, 1979-85

---

Industry	Investment	Fuels	Materials	Electricity
Metals	44.8	63.1	99.9	96.5
Textiles	41.2	51.2	99.9	97.0
Food Products	42.7	78.0	99.8	88.3
Beverages	44.0	73.9	99.8	94.1
Other Chemicals	65.3	78.4	100	96.5
Printing & Pub.	39.0	46.4	99.9	96.8
Wood Products	35.9	59.3	99.7	93.8
Apparel	35.2	34.5	99.9	97.2

---

TABLE 3  
Average Nominal Revenue Shares (Percentage), 1979-85

Industry	Unskilled	Skilled	Materials	Fuels	Electricity
Metals	15.2	8.3	44.9	1.6	1.7
Textiles	13.8	6.0	48.2	1.0	1.6
Food Products	12.1	3.5	60.3	2.1	1.3
Beverages	11.3	6.8	45.6	1.8	1.5
Other Chemicals	18.9	10.1	37.8	1.7	0.7
Printing & Pub.	19.8	10.7	40.1	0.5	1.3
Wood Products	20.6	5.3	47.0	3.0	2.4
Apparel	14.0	4.9	52.4	0.9	0.3

TABLE 4  
Unrestricted and Restricted Parameter Estimates for 8 Industries  
(Bootstrapped Standard Errors in Parentheses)

Input	Industry (ISIC Code)							
	311	381	321	331	352	322	342	313
Unskilled labor	0.138 (0.010)	0.164 (0.032)	0.138 (0.027)	0.206 (0.035)	0.137 (0.039)	0.163 (0.044)	0.192 (0.048)	0.087 (0.082)
Skilled labor	0.053 (0.008)	0.185 (0.017)	0.139 (0.030)	0.136 (0.032)	0.254 (0.036)	0.125 (0.038)	0.161 (0.036)	0.164 (0.087)
Materials	0.703 (0.013)	0.587 (0.017)	0.679 (0.019)	0.617 (0.022)	0.567 (0.045)	0.621 (0.020)	0.483 (0.028)	0.626 (0.075)
Fuels	0.023 (0.004)	0.024 (0.008)	0.041 (0.012)	0.018 (0.018)	0.004 (0.020)	0.0162 (0.016)	0.053 (0.014)	0.087 (0.027)
Capital								
unrestricted	0.13 (0.032)	0.09 (0.027)	0.08 (0.054)	0.18 (0.029)	0.17 (0.034)	0.10 (0.024)	0.21 (0.042)	0.08 (0.050)
restricted	0.14 (0.011)	0.09 (0.02)	0.06 (0.019)	0.11 (0.025)	0.15 (0.034)	0.09 (0.039)	0.21 (0.045)	0.07 (0.11)
Electricity								
unrestricted	0.038 (0.021)	0.020 (0.010)	0.017 (0.024)	0.032 (0.028)	0.017 (0.032)	0.022 (0.014)	0.020 (0.024)	0.012 (0.022)
restricted	0.011	0.015	0.014	0.021	0.005	0.008	0.011	0.012
No. Obs.	6051	1394	1129	1032	758	674	507	465

Notes: This table contains estimates from the unrestricted and restricted models.

TABLE 5

The OLS estimate minus the LP estimate

ISIC	Unskilled Labor	Skilled Labor	Materials	Fuels	Capital	Returns to Scale
311						
Mean Diff.	0.01	0.032	0.044	0.007	-0.064	0.029
% > 0	100%	100%	100%	100%	0%	100%
381						
Mean Diff.	0.024	0.003	0.022	0.006	-0.034	0.021
% > 0	100%	58.06%	100%	83.87%	3.23%	96.77%
321						
Mean Diff.	-0.004	0.014	0.013	0.002	-0.008	0.017
% > 0	29.03%	83.87%	100%	67.74%	32.26%	93.55%
331						
Mean Diff.	0.126	0.08	-0.059	0.011	-0.084	0.073
% > 0	100%	100%	3.23%	93.55%	0%	100%
352						
Mean Diff.	-0.012	0.018	0.029	-0	-0.023	0.012
% > 0	29.03%	93.55%	100%	35.48%	16.13%	64.52%
322						
Mean Diff.	0.003	0.007	0.018	0.001	-0.037	-0.008
% > 0	58.06%	54.84%	87.1%	54.84%	6.45%	38.71%
342						
Mean Diff.	-0.003	0.028	0.004	-0.003	-0.03	-0.003
% > 0	48.39%	93.55%	93.55%	25.81%	19.35%	83.87%
313						
Mean Diff.	0.083	0.001	0.042	-0.008	-0.085	0.034
% > 0	100%	64.52%	93.55%	29.03%	6.45%	77.42%

Notes: This table contains the differences between OLS and LP estimates for each coefficient across 30 bootstrapped samples. The top entry in each cell is the average difference between the OLS and LP estimate across the samples, i.e. the average across  $\beta_{OLS} - \beta_{LP}$ . The bottom entry in each cell is the percentage of samples that yielded estimates which  $\beta_{OLS} - \beta_{LP} > 0$ .

TABLE 6

Comparison of Estimators Across All 8 Industries: Testing Coefficients as a Group

Estimate of:

Significantly Different At:	$\beta_{OLS} - \beta_{LP}$	$\beta_{FE} - \beta_{LP}$	$\beta_{IV} - \beta_{LP}$	$\beta_{OP} - \beta_{LP(i>0)}$	$\beta_{LP} - \beta_{LP(i>0)}$	$\beta_{OP} - \beta_{LP}$
5%	75.0%	100%	50.0%	37.5%	50.0%	50.0%
10%	75.0%	100%	75.0%	62.5%	50.0%	62.5%
20%	100%	100%	100%	87.5%	87.5%	100%

Notes: This table provides the percent of 8 industries for which the null hypothesis that none of the 5 coefficient pairs in an industry are different is rejected in favor of the alternative that at least one of the coefficient pairs in the industry is different.

TABLE 7

Comparison of Estimators Across All 8 Industries: Individual Coefficients

Significantly Different At:	Estimate of:					
	$\beta_{OLS} - \beta_{LP}$	$\beta_{FE} - \beta_{LP}$	$\beta_{IV} - \beta_{LP}$	$\beta_{OP} - \beta_{LP(i>0)}$	$\beta_{LP} - \beta_{LP(i>0)}$	$\beta_{OP} - \beta_{LP}$
5%	32.5%	47.5%	17.5%	12.5%	15.0%	15.0%
10%	37.5%	55.0%	32.5%	20.0%	20.0%	22.5%
20%	55.0%	67.5%	55.0%	40.0%	27.5%	40.0%

Notes: This table contains the percent of significantly different coefficients between two estimators. For each pair of estimators, there are 5 coefficients which are compared at the 5%, 10%, and 20% significance levels across 8 industries, for a total of 40 coefficients; the table presents the percentage of these 40 coefficients that are significantly different for each of the three levels.

TABLE 8

Comparison of Estimators Across All 8 Industries: Returns to Scale

Estimate of:

Significantly Different At:	$\beta_{OLS} - \beta_{LP}$	$\beta_{FE} - \beta_{LP}$	$\beta_{IV} - \beta_{LP}$	$\beta_{OP} - \beta_{LP(i>0)}$	$\beta_{LP} - \beta_{LP(i>0)}$	$\beta_{OP} - \beta_{LP}$
5%	25.0%	37.5%	50.0%	12.5%	0.0%	0.0%
10%	37.5%	75.0%	50.0%	12.5%	25.0%	0.0%
20%	50.0%	75.0%	62.5%	12.5%	25.0%	12.5%

Notes: This table contains the percent of significantly different estimates of returns to scale between two estimators. For each pair of estimators, there are 8 industries which are compared at the 5%, 10%, and 20% significance; the table presents the percentage of these 8 observations on returns to scale that are significantly different for each of the three levels.



TABLE 9  
The OP estimate using minus the LP estimate using  
only firms with positive investment

ISIC	Unskilled Labor	Skilled Labor	Materials	Fuels	Capital	Returns to Scale
<b>311</b>						
Mean Diff.	-0.005	0.024	0.035	0.005	-0.042	0.017
% > 0	9.68%	100%	100%	93.55%	0%	100%
<b>381</b>						
Mean Diff.	0.021	0.006	0.018	0.003	-0.029	0.019
% > 0	90.32%	83.87%	100%	70.97%	9.68%	87.1%
<b>321</b>						
Mean Diff.	-0.006	0.02	-0.003	-0.005	0.017	0.024
% > 0	29.03%	96.77%	38.71%	22.58%	35.48%	61.29%
<b>331</b>						
Mean Diff.	0.08	0.067	-0.019	0.017	-0.09	0.055
% > 0	96.77%	100%	22.58%	83.87%	12.9%	77.42%
<b>352</b>						
Mean Diff.	-0.022	0.028	0.031	0.01	0.029	0.078
% > 0	19.35%	93.55%	96.77%	90.32%	25.81%	74.19%
<b>322</b>						
Mean Diff.	-0.034	0.003	0.009	-0	0.03	0.007
% > 0	19.35%	64.52%	61.29%	48.39%	61.29%	48.39%
<b>342</b>						
Mean Diff.	-0.029	0.007	0	-0.018	0.009	-0.03
% > 0	16.13%	51.61%	45.16%	9.68%	38.71%	22.58%
<b>313</b>						
Mean Diff.	0.057	0.025	0.055	-0	-0.028	0.108
% > 0	83.87%	77.42%	90.32%	48.39%	32.26%	64.52%

Notes: This table contains the differences between OP coefficient estimates and the LP coefficient estimates from firms just reporting positive investment. The top entry in each cell is the average difference between the estimates across 30 bootstrapped samples. The bottom entry in each cell is the percentage of samples that yielded estimates which  $\beta_{OP} - \beta_{LP}(i > 0) > 0$ .

TABLE 10

Comparison of Estimates Across the Electricity, Materials, and Fuels Proxies  
Standard Errors in Parentheses

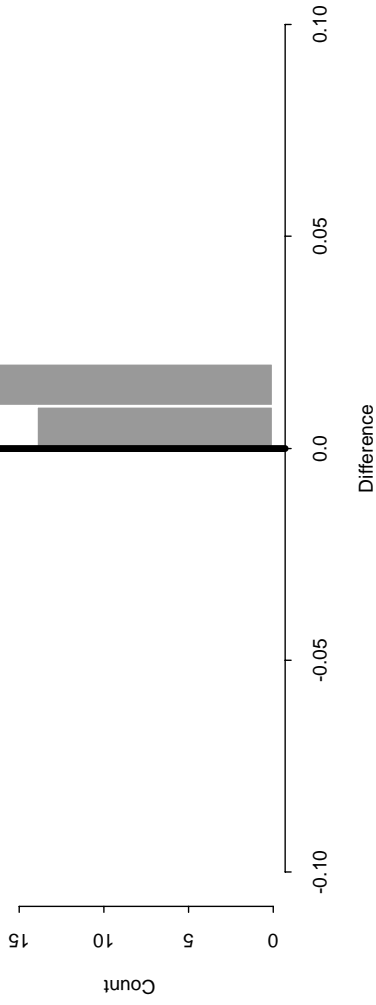
Coefficient	ISIC 311			ISIC 381		
	Electricity	Materials	Fuels	Electricity	Materials	Fuels
Unskilled Labor	0.140 (0.014)	0.137 (0.014)	0.136 (0.014)	0.177 (0.031)	0.176 (0.031)	0.173 (0.031)
Skilled Labor	0.055 (0.003)	0.052 (0.003)	0.057 (0.004)	0.180 (0.02)	0.182 (0.021)	0.187 (0.022)
Capital	0.15 (0.014)	0.15 (0.008)	0.16 (0.013)	0.12 (0.018)	0.21 (0.023)	0.09 (0.018)

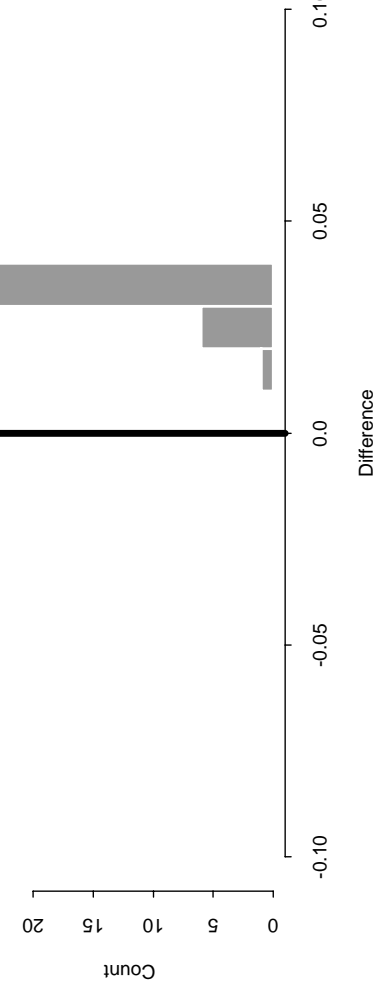
Coefficient	ISIC 321			ISIC 331		
	Electricity	Materials	Fuels	Electricity	Materials	Fuels
Unskilled Labor	0.152 (0.028)	0.159 (0.029)	0.146 (0.028)	0.216 (0.036)	0.207 (0.033)	0.198 (0.030)
Skilled Labor	0.135 (0.024)	0.150 (0.024)	0.148 (0.027)	0.134 (0.035)	0.136 (0.035)	0.134 (0.034)
Capital	0.18 (0.261)	0.17 (0.027)	0.16 (0.311)	0.10 (0.017)	0.09 (0.035)	0.06 (0.023)

Notes: This table compares the point estimates and standard errors across three different proxy variables: electricity, materials, and fuels.

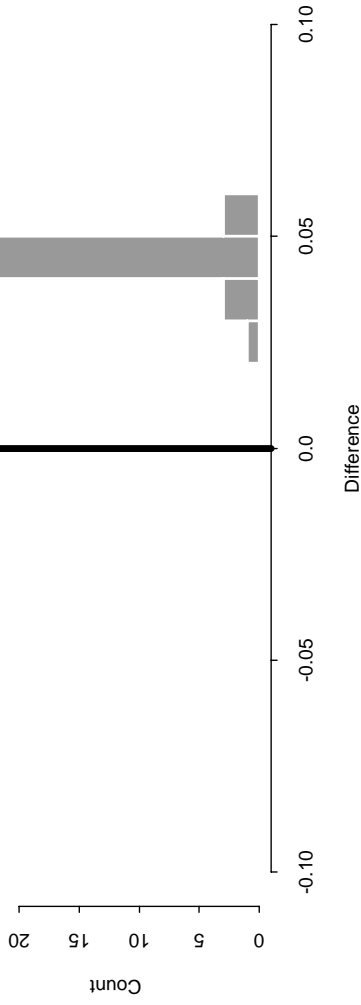
Histogram of differences between OLS and LP for unskilled labor



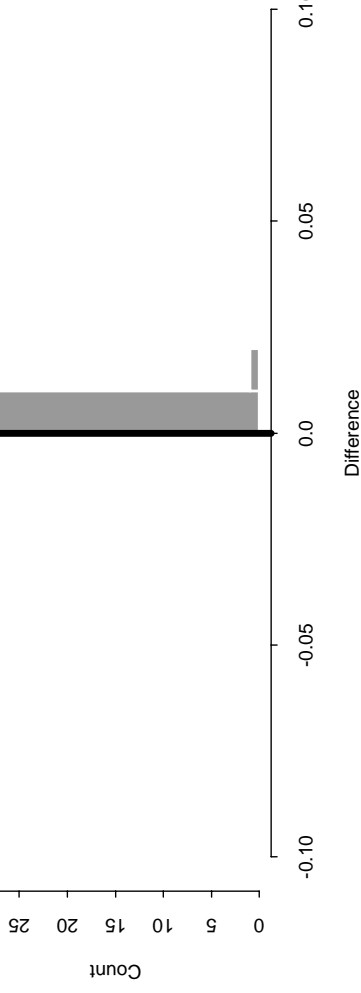
Histogram of differences between OLS and LP for skilled labor



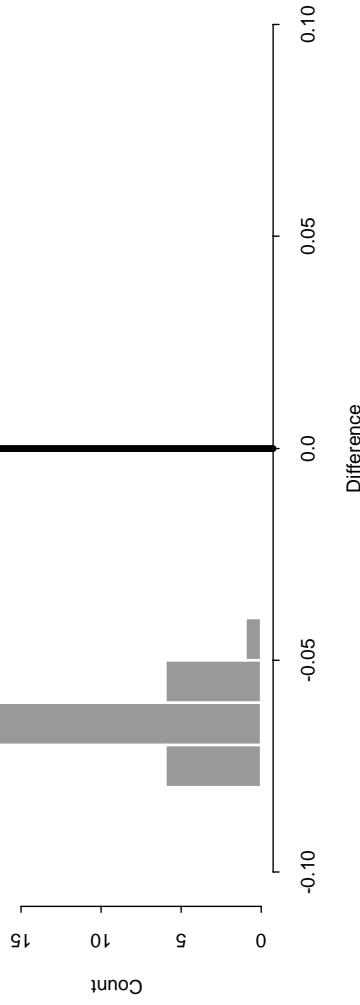
Histogram of differences between OLS and LP for materials



Histogram of differences between OLS and LP for fuels



Histogram of differences between OLS and LP for capital



Histogram of differences between OLS and LP for returns for scale

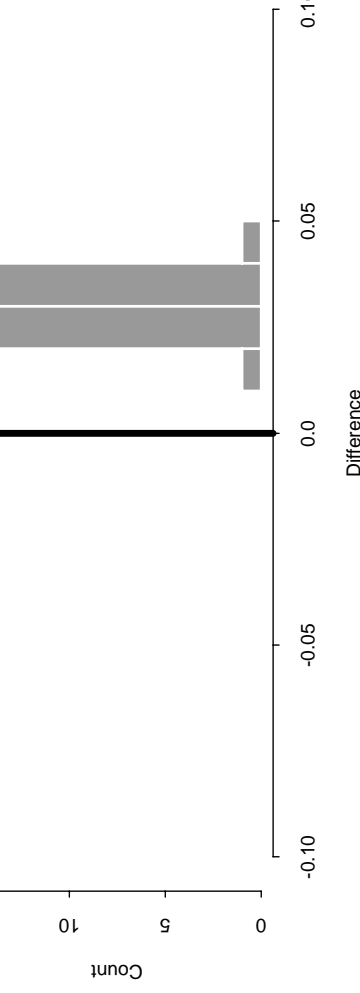


TABLE A1

The FE estimate minus the LP estimate

ISIC	Skilled Labor	Unskilled Labor	Materials	Fuels	Capital	Returns to Scale
311						
Mean Diff.	0.045	-0.036	-0.063	0.008	-0.116	-0.162
% > 0	100%	0%	0%	93.55%	0%	0%
381						
Mean Diff.	0.078	-0.099	-0.087	0.013	-0.035	-0.131
% > 0	96.77%	0%	0%	74.19%	12.9%	3.23%
321						
Mean Diff.	0.05	-0.11	-0.107	0.045	0.039	-0.082
% > 0	87.1%	3.23%	0%	100%	90.32%	3.23%
331						
Mean Diff.	0.083	0.02	-0.205	0.007	-0.08	-0.177
% > 0	83.87%	61.29%	0%	61.29%	6.45%	0%
352						
Mean Diff.	0.169	-0.086	-0.145	0.023	-0.027	-0.066
% > 0	100%	3.23%	0%	87.1%	29.03%	16.13%
322						
Mean Diff.	0.047	-0.11	-0.129	0.028	-0.04	-0.203
% > 0	80.65%	0%	0%	87.1%	35.48%	3.23%
342						
Mean Diff.	0.208	-0.093	-0.283	0.021	-0.093	-0.24
% > 0	100%	6.45%	0%	80.65%	0%	0%
313						
Mean Diff.	0.261	-0.088	-0.14	-0.076	-0.105	-0.148
% > 0	100%	12.9%	0%	0%	9.68%	19.35%

Notes: This table contains the differences between FE and LP estimates for each coefficient across 30 bootstrapped samples. The top entry in each cell is the average difference between the FE and LP estimate across the samples, i.e. the average across  $\beta_{FE} - \beta_{LP}$ . The bottom entry in each cell is the percentage of samples that yielded estimates which  $\beta_{FE} - \beta_{LP} > 0$ .

TABLE A2

The IV estimate minus the LP estimate

ISIC	Unskilled Labor	Skilled Labor	Materials	Fuels	Capital	Returns to Scale
311						
Mean Diff.	-0.005	0.023	0.078	0.012	-0.077	0.03
% > 0	19.35%	100%	100%	93.55%	0%	100%
381						
Mean Diff.	-0.003	0.016	0.045	-0.001	-0.039	0.018
% > 0	51.61%	67.74%	96.77%	41.94%	6.45%	90.32%
321						
Mean Diff.	-0.04	0.046	0.04	-0.013	-0.007	0.026
% > 0	6.45%	96.77%	96.77%	3.23%	29.03%	100%
331						
Mean Diff.	0.127	0.073	-0.051	0.021	-0.082	0.087
% > 0	93.55%	93.55%	9.68%	83.87%	0%	100%
352						
Mean Diff.	-0.054	-0.003	0.097	-0.025	-0.014	-0.001
% > 0	3.23%	41.94%	100%	3.23%	22.58%	45.16%
322						
Mean Diff.	-0.186	0.026	0.064	0.045	-0.006	-0.057
% > 0	0%	83.87%	93.55%	100%	51.61%	0%
342						
Mean Diff.	-0.028	0.027	0.072	-0.013	-0.029	0.028
% > 0	16.13%	80.65%	90.32%	22.58%	16.13%	80.65%
313						
Mean Diff.	0.052	-0.039	0.139	-0.039	-0.098	0.015
% > 0	67.74%	35.48%	83.87%	22.58%	9.68%	61.29%

Notes: This table contains the differences between IV and LP estimates for each coefficient across 30 bootstrapped samples. The top entry in each cell is the average difference between the IV and LP estimate across the samples, i.e. the average across  $\beta_{IV} - \beta_{LP}$ . The bottom entry in each cell is the percentage of samples that yielded estimates which  $\beta_{IV} - \beta_{LP} > 0$ .