



OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

The Review of Economic Studies, Ltd.

Estimating Production Functions Using Inputs to Control for Unobservables

Author(s): James Levinsohn and Amil Petrin

Source: *The Review of Economic Studies*, Vol. 70, No. 2 (Apr., 2003), pp. 317-341

Published by: [Oxford University Press](#)

Stable URL: <http://www.jstor.org/stable/3648636>

Accessed: 06/08/2013 15:49

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Oxford University Press and *The Review of Economic Studies, Ltd.* are collaborating with JSTOR to digitize, preserve and extend access to *The Review of Economic Studies*.

<http://www.jstor.org>

Estimating Production Functions Using Inputs to Control for Unobservables

JAMES LEVINSOHN

University of Michigan

and

AMIL PETRIN

University of Chicago

First version received June 2000; final version accepted October 2002 (Eds.)

We add to the methods for conditioning out serially correlated unobserved shocks to the production technology. We build on ideas first developed in Olley and Pakes (1996). They show how to use investment to control for correlation between input levels and the unobserved firm-specific productivity process. We show that intermediate inputs (those inputs which are typically subtracted out in a value-added production function) can also solve this simultaneity problem. We discuss some theoretical benefits of extending the proxy choice set in this direction and our empirical results suggest these benefits can be important.

1. INTRODUCTION

Economists began relating output to inputs in the early 1800's. A large literature on estimating production functions has followed, in part because much of economic theory yields testable implications that are related to the technology and optimizing behaviour.¹

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 productivity 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 (OLS) estimates of production functions will yield biased parameter estimates, and, by implication, biased estimates of productivity.

Many alternatives to OLS have been proposed, and we add to this set by extending Olley and Pakes (1996). They show the conditions under which an investment proxy controls for correlation between input levels and the unobserved productivity shock. Their approach has the advantage that, for many questions, it is no more difficult to implement than OLS. We show when intermediate inputs (those inputs which are typically subtracted out in a value-added production function) can also solve this simultaneity problem. We discuss some potential benefits of expanding the choice set of proxies to include these inputs.

1. Much of the early applied work exploring this relationship was pioneered by agricultural economists like Von Thunen (a colleague of Cournot), who collected data at his farm in the 1820's to measure the marginal product of inputs and the substitutability between inputs. Flux (1913), using one of the first available manufacturing censuses, details relationships between inputs and output for manufacturing firms in England. Chambers (1997) provides a brief history of production function estimation.

One benefit is strictly data-driven. It turns out that the investment proxy is only valid for plants reporting non-zero investment. (This is due to an invertibility condition described below.) Pronounced adjustment costs, which do not necessarily 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 truncating all of these plants. Using intermediate input proxies instead of investment avoids this problem. This is because firms (in our data, at least) almost always report positive use of intermediate inputs like materials or electricity.²

To the extent that non-convex adjustment costs are an important issue, intermediate inputs may confer another benefit. If adjustment costs lead to kink points in the investment demand function, plants may not entirely respond to some productivity shocks, and correlation between the regressors and the error term can remain. If it is less costly to adjust the intermediate input, it may respond more fully to the entire productivity term.

Another nice feature of the intermediate input is that it provides a simple link between the estimation strategy and the economic theory, primarily because intermediate inputs are not typically state variables. We develop this link, deriving the conditions that must hold if intermediate inputs are to be a valid proxy for the productivity shock. We suggest three specification tests for evaluating any proxy's performance and for comparing among proxies when more than one is available. We also derive the expected directions of bias on the OLS estimates relative to our intermediate input approach when simultaneity exists. We take the framework to four Chilean manufacturing industries, finding significant differences between OLS and our approach that are also consistent with simultaneity.³

Many estimators have been developed to address simultaneity under different data generating processes to which OLS is not robust. We compare estimates between OLS, fixed effects, the Olley–Pakes investment proxy estimator, our intermediate input proxy estimator and a Blundell–Bond GMM estimator (a lagged-input instrumental variables (IV) estimator with fixed effects, time effects, AR(1) and MA(0) shocks, all of which raise potential simultaneity problems). While these models do not generally nest one another, any test rejecting no differences between estimates tells us that the two processes cannot both be compatible with the industry under consideration. These results add to the evidence of a simultaneity problem and shed some light on its underlying nature.

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. Appendices include a monotonicity proof, a short-cut for simplifying estimation, and a “recipe” for our estimation routine.

2. ESTIMATION IN THE PRESENCE OF SIMULTANEITY

We write firm i 's production at time t as $y_{it} = f(x_{it}, \epsilon_{it}; \beta)$ with β parameters. x_{it} includes inputs that are easily adjusted and those that evolve over time in response to beliefs. The errors $\{\epsilon_{it}\}_{t=1}^{\infty}$ are often thought of as Hicks neutral productivity shocks.

2. For many firm-level data sets this truncation is not trivial. Although researchers working with the longitudinal research database from the U.S. (as did Olley and Pakes) or comparable manufacturing censuses from the U.K. or France may not find half of their sample reporting zero investment, much of the plant-level research being conducted today is on easier-to-obtain data from countries like Turkey, Columbia, Mexico and Indonesia. In these countries, as well as Chile, the “zero investment” problem is more likely to loom large.

3. In this paper, we only report the results from the four largest manufacturing industries. In earlier versions of the paper, we report results for eight industries. This is an effort to keep the number of reported results manageable. Readers interested in seeing results for all of the industries can access the NBER Working Paper 6893 (Levinsohn and Petrin, 1999).

A simultaneity problem arises when there is contemporaneous correlation between x_{it} and ϵ_{it} . This simultaneity violates the OLS conditions for unbiased and consistent estimation. It can arise with firm-level data when input choices respond to shocks. Marschak and Andrews (1944) suggest this problem may be most pronounced for inputs that adjust rapidly. Applied researchers have spent much effort addressing the econometric problem these correlations confer.

In a multivariate context it is generally impossible to sign the biases of the OLS coefficients when simultaneity exists and there are many inputs. Some intuition about the bias can be derived from an analysis of the OLS estimates for a two-input production function, with one freely variable input l_{it} (call it labour) and one quasi-fixed input k_{it} (call it capital):

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

OLS estimates for the inputs are

$$\hat{\beta}_l = \beta_l + \frac{\hat{\sigma}_{k,k}\hat{\sigma}_{l,\epsilon} - \hat{\sigma}_{l,k}\hat{\sigma}_{k,\epsilon}}{\hat{\sigma}_{l,l}\hat{\sigma}_{k,k} - \hat{\sigma}_{l,k}^2},$$

and, symmetrically,

$$\hat{\beta}_k = \beta_k + \frac{\hat{\sigma}_{l,l}\hat{\sigma}_{k,\epsilon} - \hat{\sigma}_{l,k}\hat{\sigma}_{l,\epsilon}}{\hat{\sigma}_{l,l}\hat{\sigma}_{k,k} - \hat{\sigma}_{l,k}^2},$$

where $\hat{\sigma}_{a,b}$ denotes the sample covariance between a and b .

We consider bias in three different cases. Since the denominator is always positive, the sign of the bias is determined by the numerator. If only labour responds to the shock (say more labour is hired in response to a productivity shock, so $\sigma_{l,\epsilon} > 0$), and capital is not correlated with labour, then $\hat{\beta}_l$ will tend to be biased up but $\hat{\beta}_k$ will remain unbiased. If only labour responds to the shock and capital and labour are positively correlated a negative bias on the capital coefficient can also result. Finally, if capital and labour are positively correlated and labour's correlation with the productivity shock is higher than capital's correlation, $\hat{\beta}_l$ will tend to overestimate β_l and $\hat{\beta}_k$ will usually underestimate β_k . For short panels we think these last two cases may be most relevant, as between-firm variation often plays a dominant role in identification, and capital and labour tend to be highly correlated in this dimension.

Within estimators are a common alternative to OLS, using only the variation within-firm to protect against potential correlation between unobserved firm-specific fixed effects (like managerial quality) and input choices. Sometimes, the between-firm variation can be important for obtaining precise estimates of output elasticities associated with state variables (in short panels, firms may not adjust capital much). Thus within estimators offer more protection against firm-specific effects than OLS, but they can exacerbate other problems by reducing the "signal".

An instrumental variable (IV) estimator achieves consistency by instrumenting the explanatory variables with regressors that are correlated with the inputs but uncorrelated with ϵ_{it} . The IV approach can also alleviate measurement error problems, which tend to be most pronounced in capital.⁴ Potential instruments at the firm-level include input prices and lagged values of input use. Firm-level input prices are rarely observed. Lagged values of inputs are valid instruments if the lag time is long enough to break the dependence between the input choices and the serially correlated shock. Blundell and Bond (2000) develop a sophisticated cousin of the IV

4. Measurement error in capital has the same implications for the parameter estimates as the simultaneity problem described above. In particular, if capital and labour are positively correlated and capital is measured with error, the noise will tend to attenuate capital's coefficient towards zero and its associated output change will be incorrectly attributed to labour.

approach that is robust to firm-specific fixed effects, serially correlated productivity shocks and measurement error.⁵

The investment proxy

Olley and Pakes (1996) suggest a novel approach to addressing this simultaneity problem. They include in the estimation equation a proxy which they derive from a structural model of the optimizing firm. The proxy controls for the part of the error correlated with inputs by “annihilating” any variation that is possibly related to the productivity term.

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

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \omega_t + \eta_t. \quad (1)$$

Inputs are divided into a freely variable one (l_t) and the state variable capital (k_t).⁶ ϵ_t is assumed to be additively separable in a transmitted component (ω_t) and an i.i.d. component (η_t). The key difference between ω_t and η_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.

Olley–Pakes write investment as just a function of the two state variables in their model, k_t and ω_t , or

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

When ω_t is stochastically increasing in past values, Pakes (1996) proves that optimizing firms *choosing to invest* have investment functions that are strictly increasing in the unobserved productivity shock. Basically, better productivity shocks today mean better shocks in the future, and this leads to capital accumulation.

The monotonicity allows $i_t(\omega_t, k_t)$ to be inverted to yield ω_t as a function of investment and capital, or $\omega_t = \omega_t(i_t, k_t)$. One can then rewrite (1) as

$$y_t = \beta_l l_t + \phi_t(i_t, k_t) + \eta_t, \quad (2)$$

where

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

A first-stage estimator that is linear in l_t and non-parametric in ϕ_t can be used to obtain a consistent estimate of β_l .⁷ Olley and Pakes use a fourth-order polynomial in i_t and k_t to approximate $\phi(\cdot)$, estimating (2) using OLS, with output regressed on labour and the polynomial terms.

We follow the exposition in Robinson (1988) to illustrate the idea further. It is suggestive of how one might implement alternative non-parametric estimators (which we do in what follows). Robinson (1988) takes the expectation of equation (2) conditional on i_t and k_t . This is given by

$$E[y_t | i_t, k_t] = \beta_l E[l_t | i_t, k_t] + \phi_t(i_t, k_t) \quad (3)$$

5. 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 $\{X_{it}, \epsilon_{it}\}_{t=1}^{\infty}$ 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 ϵ_{it} and X_{it} over time because the process $\{X_{it}, \epsilon_{it}\}_{t=1}^{\infty}$ follows a path that depends upon its starting values (X_{i1}, ϵ_{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 $\{X_{it}, \epsilon_{it}\}_{t=1}^{T-1}$ is independent of $\{X_{it}, \epsilon_{it}\}_{t=T}^{\infty}$, where T is the first date a firm is observed. A second solution splits the sample into two parts, the first part of which is used to estimate starting values (see Roberts and Tybout, 1997).

6. For simplicity, we assume (as they do) that capital is the only state variable over which the firm has control.

7. We will always use $\phi_t(\cdot)$ when discussing the non-parametric part of this first stage; its 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.

because: (i) η_t is mean independent of i_t and k_t ; and (ii) $E[\phi_t(i_t, k_t) \mid i_t, k_t] = \phi_t(i_t, k_t)$. Subtracting equation (3) from (2) yields

$$y_t - E[y_t \mid i_t, k_t] = \beta_l(l_t - E[l_t \mid i_t, k_t]) + \eta_t. \tag{4}$$

By assumption η_t is mean independent of l_t (and thus of the transformed regressor $l_t - E[l_t \mid i_t, k_t]$), so no-intercept OLS can be used to obtain consistent estimates of β_l .

Since capital enters $\phi(\cdot)$ twice, a more complete model is needed to identify β_k . Olley and Pakes assume that ω_t follows a first-order Markov process and that capital does not immediately respond to ξ_t , the innovations in productivity over last period's expectation, given by

$$\xi_t = \omega_t - E[\omega_t \mid \omega_{t-1}].$$

Defining y_t^* as output net of labour's contribution, they write

$$y_t^* = y_t - \beta_l l_t = \beta_0 + \beta_k k_t + E[\omega_t \mid \omega_{t-1}] + \eta_t^*, \tag{5}$$

where $\eta_t^* = \xi_t + \eta_t$. Under these assumptions regressing y_t^* on k_t and a consistent estimate of $E[\omega_t \mid \omega_{t-1}]$ produces a consistent estimate of β_k (because both ξ_t and η_t are uncorrelated with k_t).⁸ When this approach works, it can have advantages relative to OLS, within, and traditional instrumental variable estimators (see Griliches and Mairesse, 1998).⁹

When the investment proxy might fail

Investment is a control on a state variable, something which by definition is costly to adjust. Costs of adjustment can cause problems for estimation in different ways. Firms that make only intermittent investments will have their zero-investment observations truncated from the estimation routine (the monotonicity condition does not hold for these observations). For manufacturing censuses this can be a large portion of the data.

While this truncation issue relates only to efficiency, non-convex adjustment costs may lead to kinks in the investment function that affect the responsiveness of investment to the transmitted shock even when investment is undertaken.¹⁰ For example, suppose $i_t(\omega_t, k_t)$ has some maximal level of investment for all possible outcomes of ω_t . Then $i_t(\omega_t, k_t) = \bar{i}_t(\omega_t, k_t)$ when $\omega_t \geq \bar{\omega}_t(k_t)$, for the kink point $\bar{\omega}_t(k_t)$. The error term in (4) becomes $\eta_t + (\omega_t - \bar{\omega}_t(k_t))$, which is correlated with l_t . Alternatively, suppose $\bar{\omega}_t$ is instead the extent to which ω_t is known at the time of the investment decision, and that $i_t = i_t(\bar{\omega}_t, k_t)$ is monotonic in $\bar{\omega}_t$. Again, $(\omega_t - \bar{\omega}_t)$ remains in the error term. Of course in both cases the investment proxy is helpful because it controls for $\bar{\omega}_t$.

8. Note that β_0 is not separately identified from the mean of $E[\omega_t \mid \omega_{t-1}]$ without some further restriction.
 9. From Griliches and Mairesse (1998) with the variable references changed 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 ω , the unobserved transmitted component of ϵ . Trying to proxy for the unobserved ω (if it can be done correctly) has several advantages over the usual within estimators (or the more general Chamberlain and GMM type estimators): it does not assume that ω reduces to a "fixed" (over time) firm effect; it leaves more identifying variance in l 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.

10. Doms and Dunne (1998) and Attanasio, Pacelli and Dos Reis (2000) report lumpy inertial behaviour in investment data from U.S. and U.K. plant-level surveys, suggesting non-convex adjustment costs exist.

3. INTERMEDIATE INPUTS AS PROXIES

We now add a second freely variable input, ι , which we call the intermediate input (perhaps materials or energy).¹¹ Writing the log of output as a function of the log of inputs and the shocks we have

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \beta_\iota \iota_t + \omega_t + \eta_t. \quad (6)$$

The intermediate input's demand function is given as

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

and it must be monotonic in ω_t for all (relevant) k_t to qualify as a valid proxy. Note that input and output prices are assumed to be common across firms (they are suppressed), and there is no error in the input demand function.¹² In the data section we use these conditions to help choose between candidate proxies.

Assuming monotonicity holds one can invert the input demand function to obtain $\omega_t = \omega_t(\iota_t, k_t)$. Thus, the intermediate input replaces investment, with $\phi_t(\cdot)$ now given as a function of the intermediate input and capital, or

$$\phi_t(\iota_t, k_t) = \beta_0 + \beta_k k_t + \beta_\iota \iota_t + \omega_t(\iota_t, k_t). \quad (7)$$

The equation for the second stage changes to

$$y_t^* = \beta_0 + \beta_k k_t + \beta_\iota \iota_t + E[\omega_t | \omega_{t-1}] + \eta_t^*. \quad (8)$$

Similar to investment, for any value of (β_k, β_ι) we can estimate $E[\omega_t | \omega_{t-1}]$. While $E[k_t \eta_t^*] = 0$ is assumed to still hold for (8), $E[\iota_t \eta_t^*] = 0$ does *not* generally hold because the intermediate input is correlated with η_t^* (it responds to ξ_t). Since firms choose ι_{t-1} before either component of η_t^* is realized, it should be uncorrelated with η_t^* . It should also be correlated with ι_t (via, for example, size correlation over time due to irreversibility in capital investment and/or the persistence in ω_t), so we use $E[\iota_{t-1} \eta_t^*] = 0$ to obtain identification.

The monotonicity condition

The monotonicity condition for intermediate inputs is identical to that for investment; conditional on capital, profit maximizing behaviour must lead more productive firms to use more intermediate inputs. We imagine a story where increases in productivity increase the intermediate input's marginal productivity. This in turn leads firms to increase output, which leads to more input use.

Aside from some regularity conditions on the production function $f(\cdot)$, conditional on k , the sign of the change in intermediate input use for a small change in ω is given by

$$\text{sign}\left(\frac{\partial \iota}{\partial \omega}\right) = \text{sign}(f_{ll} f_{l\omega} - f_{ll} f_{l\omega}) \quad (9)$$

(where f_{ll} is the second derivative of $f(\cdot)$ with respect to l , etc.).¹³ Optimizing behaviour implies the marginal product of labour declines as labour increases, so $f_{ll} < 0$ for chosen input bundles. If increases in productivity always weakly increase the marginal product of inputs, then $f_{l\omega} \geq 0$ (and $f_{l\omega} \geq 0$), so $-f_{ll} f_{l\omega} \geq 0$. If $f_{ll} = 0$ or $f_{l\omega} = 0$ and $f_{l\omega} > 0$ the monotonicity condition

11. The contribution of these inputs is usually subtracted from output before estimation (implicitly) using a separability assumption. Thus, their use adds little burden to necessary data requirements already faced when estimating production functions.

12. Loosening the latter assumption is the subject of ongoing work.

13. Appendix A contains the full derivation.

holds. If $f_{l\omega} > 0$ the result is driven by the cross-partial of output with respect to the intermediate input and labour. If the marginal product of the intermediate input weakly increases as labour use increases ($f_{li} \geq 0$) then the result holds. Even if the marginal product falls with increases in labour, the condition may still hold (it will depend on relative magnitudes of the two products on the R.H.S. of (9)).

One advantage of this estimator is that it is easy to verify whether the monotonicity condition is consistent with some common technologies used by economists (e.g. Cobb–Douglas or constant elasticity of substitution). 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 using the firm's dynamic problem, and this can be difficult (as Pakes demonstrates). Another convenient property of this estimator is that the monotonicity condition is not imposed, so we can check to see if it is empirically justified. We show how to do this in Section 6.

Common input prices (and other common unobserved factors)

As with investment, the intermediate input demand equation $\iota_t = \iota_t(\omega_t, k_t)$ is not indexed by other factors (like input prices). If input prices are observed and not common across firms, they can be included directly in the demand function, loosening the common factor price restriction. If they are not observed but one suspects input price ratios vary over time, or by region, or by urban/rural location, or by type of firm, one can estimate different functions for these time periods or regions (if they are observed), making estimation robust to these differences at the expense of placing greater demands on the data.

4. DATA

In order to implement the intermediate input proxy, we need data. We use an 8-year panel from Chile that has also been used elsewhere.¹⁴ This Chilean data is representative of many firm-level panels in the sense that it has many firm-level variables (including many intermediate inputs), it is not censored for entry and exit, and it has a reasonable time-series dimension to it.

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 10 employees and collected by Chile's Instituto Nacional de Estadística (INE). A very detailed description of how the longitudinal samples were combined into a panel is found in Lui (1991).¹⁵

In an attempt to keep the analysis manageable, we focus on the four largest industries (excluding petroleum and refining). The 3-digit level industries (along with their ISIC codes) are Metals (381), Textiles (321), Food Products (311) and Wood Products (331).¹⁶ The data are observed annually and they include gross revenue (our output index), indices of labour and capital inputs, and a measure of the intermediate inputs electricity, materials, and fuels.¹⁷ Labour

14. See, for example, Tybout, De Melo and Corbo (1991), Lui (1993), Lui and Tybout (1996), Levinsohn (1999) and Pavcnik (1999).

15. 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).

16. Results for the next four largest industries, Other Chemicals (352), Beverages (313), Printing and Publishing (342) and Apparel (322) are reported in NBER Working Paper 7819 and also available at the websites of both authors.

17. Revenue is our measure of plant output because (as in most firm-level data) we do not observe a plant-level measure of physical output. Many examples of output being directly observed are found in the voluminous literature in agricultural economics.

is the number of man-years hired for production, and firms distinguish between their blue- and white-collar workers. Gross revenue, capital, materials, electricity, and fuels each have their own annual price deflator (most of them provided by the Banco Central de Chile) and are each deflated to real 1980 Chilean pesos.

Construction of the capital variable is documented in Lui (1991). It is the sum of the real peso value of depreciated buildings, machinery and vehicles, each of which is assumed to have a depreciation rate (δ) of 5, 10 and 20% respectively.¹⁸ Thus each type of capital K_j evolves according to

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

and the total capital index at time t is

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

Our capital variable is constructed in a slightly different manner from Olley–Pakes as they assume investment reported last period enters the production function as capital in this period. We assume investment occurring in this period enters capital in this period. Obviously, the details on the timing of data collection, the timing of the actual investment decision, and the capital adjustment process will determine which (if either) of these assumptions is appropriate. We do not know these details for our data, so for us the choice is somewhat arbitrary, although the decision affects the proxy's implementation.¹⁹ Under our accumulation process, today's investment decision must be made knowing *only* the outcome of ω_{t-1} , or capital (via investment) will respond to ω_t , violating the consistency condition.²⁰ Under this scenario, *next* period's investment is the proxy for this period's shock (it responds fully to ω_t). In Olley–Pakes this period's investment is the proxy for this period's ω_t , and last period's investment enters capital this period. The subtleties of timing are clearly important for these variants of the proxy approach.

Table 1 provides some macroeconomic background as well as some summary statistics for the industries we examine. By 1979, most of Pinochet's economic policies were already in place. The Latin American debt crisis led to a recession in 1982 and 1983 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 1979–1986. We will take these macroeconomic cycles into account by allowing the $\omega_t(\iota_t, k_t)$ function to be different for each of these three different time periods (so $t = 1, 2, 3$).

It is also evident from Table 1 that this period is characterized by major consolidation and exit; the number of plants falls in every industry from the beginning to the end of the sample (although there is also entry in our sample). 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). They also show once they move to the unbalanced panel, their selection correction does not change their results. We simply note that our sample is unbalanced, and we do not focus on selection issues.²¹

18. No initial capital stock is reported for some plants, although investment is recorded. When possible, we used a capital series that was reported for a subsequent base year. For a small number of plants, capital stock is not reported in any year. For these plants, we estimated a projected initial capital stock based on other reported plant observables. We then used the investment data to fill out the capital stock data.

19. Not knowing the timing of the actual investment decisions and capital adjustment process is the current norm for econometric research because the details of investment behaviour are almost never reported.

20. This story is consistent if, for example, investment recorded today was ordered before this period's shock is known (say at the end of last period) and it comes on-line this period.

21. We have experimented with dual index selection corrections of Olley and Pakes and found as they and Griliches and Mairesse (1998) did that the selection criterion made little difference once the simultaneity correction was in place. In order to focus on the intermediate inputs issue, we do not include those methods or results in this paper.

TABLE 1
Some descriptive statistics on Chilean manufacturing

Year	GDP	Industry (ISIC)							
		Food products (311)		Metals (381)		Textiles (321)		Wood products (331)	
		Plants	Value added	Plants	Value added	Plants	Value added	Plants	Value added
1979	997.6	1537	39.0	459	10.0	503	12.4	524	10.4
1980	1075.3	1439	43.4	447	11.0	445	12.9	449	8.7
1981	1134.7	1351	42.7	413	11.5	403	11.3	406	6.8
1982	974.9	1319	47.0	365	8.1	350	8.7	358	6.5
1983	968.0	1297	42.9	322	8.3	327	9.7	335	8.1
1984	1029.4	1340	46.8	358	11.4	336	10.4	339	10.3
1985	1054.6	1338	49.1	351	9.6	337	10.8	342	10.1
1986	1114.3	1288	61.4	347	9.6	331	12.9	313	5.3

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

TABLE 2
Per cent of non-zero observations

Industry (ISIC)	Investment	Fuels	Materials	Electricity
Food products (311)	42.7	78.0	99.8	88.3
Metals (381)	44.8	63.1	99.9	96.5
Textiles (321)	41.2	51.2	99.9	97.0
Wood products (331)	35.9	59.3	99.7	93.8

Choosing among intermediate inputs

The discussion thus far has focused on using intermediate inputs as the proxy variable. In practice, there are typically several intermediate inputs and the question of choosing among them naturally arises. Details of the industry and the frequency and type of questions asked of firms can play an important role in choosing among inputs. One natural way to start evaluating the potential usefulness of a proxy is to count its zero values. In general, the number of zeros bounds from below the number of observations that must be truncated from the estimation routine. Table 2 lists the percentage of firm-level observations reporting non-zero levels of investment, fuels, materials, and electricity. It suggests that there is significant variability in zero vs. non-zero use across inputs. As described earlier, these zero observations may also reflect kinks in the factor demand curves arising from (for example) adjustment costs, which can violate the monotonicity condition.

Table 2 indicates that, in our data, many firms do not undertake investment every period (year). For these observations, no Olley–Pakes proxy is available. This leads to a rather severe efficiency loss as we would have to truncate over 50% of the observations in each industry. For Olley–Pakes, who use data on larger firms in the capital-intensive U.S. telecommunications industry, only 8% of firm/year observations are zero.

Positive use of materials is reported for over 99% of the sample's firm-year observations for all four industries. For electricity the fraction of non-zero observations is only slightly lower. Fuels are non-zero for most observations, but materials and/or electricity are preferred as a proxy by this "non-zeros" criterion. We will focus principally on these two candidates, although the ideas we discuss apply broadly to any input under consideration.

Except for electricity, we have very little information on input prices. Electricity prices were fully regulated by 1982, with the law requiring that all plants consuming less than 2 MW

of electricity be able to purchase electricity at a fixed and common unit price. This covers over 90% of our sample.²²

Two further considerations arising from the estimation assumptions can help provide guidance in choosing among inputs. First, the estimation approach assumes there is no error in the input demand equation, so for any capital level and productivity shock, a firm is assumed to readily be able to obtain $l_t(\omega_t, k_t)$. This may be problematic for electricity, which in Chile at the time was not very reliably generated or delivered. Plant slowdowns and shutdowns caused by unreliable supply may lead to observed electricity usage that is different from true demand. To some extent, a similar story for materials and/or fuels may hold, especially for firms located in areas where events like bad weather can lead to disruptions in delivery.

Second, while measurement error is always a concern, econometric theory tells us that it takes on a heightened importance when using non-parametric estimators. Hence inputs measured with less error are generally preferred as proxies. On this matter, potential measurement problems arise if inputs are stored period to period and changes in inventories of inputs are not directly observed (for example, firms only report new input purchases).²³ In our data firms record the amount of electricity they purchase, generate, and sell, so we can compute consumption directly. The inability to store electricity for long periods means that its use should be highly correlated with the year-to-year productivity terms. Materials and fuels, on the other hand, may be easy to store over time, and hence new purchases of these inputs (which we observe) may not exactly track inputs used in production.

Three specification tests

Because the choice of a proxy has an arbitrary (albeit informed) element, we suggest three specification tests. First, an informal but important specification check is to plot the proxy as a function of its two explanatory variables. To be empirically consistent with the model, the productivity shock should increase in the use of the intermediate input, holding the capital level constant. If the function is monotonic but *decreasing*, or if the function does not satisfy monotonicity, one might need to group firms according to some other observable(s) to loosen the common factor price restriction. One could still use the proxy in principle when the function is monotone decreasing conditional on capital, but the interpretation of the “productivity” term that it’s proxying for must be modified in a way that makes the theory consistent with this result (*i.e.* why does it decrease as input use increases, conditional on capital?).

A second test asks whether we get the same estimates using either electricity or materials. Not rejecting that the estimates are the same suggests either input and the single factor (ω) may be sufficient for modelling production. Unfortunately, this test does not provide clear guidance as to the problem if one rejects; rejection does not mean that both model specifications fail (one may be correct).

Finally, as suggested in Olley–Pakes, the freely variable input, labour, chosen in this period should not be correlated with the innovation in productivity next period (*i.e.* $\text{Corr}(l_t, \xi_{t+1}) = 0$). We extend this test to include all six inputs, and this provides us with six over-identifying conditions that we use to test the framework.

22. The law requires the regulated price be within a 10% band around the average price in the freely negotiated contracts (everyone over 2 MW). For a discussion of these and other issues relating to the regulation of electricity in Chile, see Bitran and Saez (1994).

23. Input inventories may be likely to occur when the storage costs are low and delivery is not just-in-time or input prices vary significantly over time.

5. ESTIMATION

In this section, we cover only the specifics of how our estimation routine is implemented. Consistency proofs for our estimators would use results from Pakes and Olley (1995). Readers interested in implementing our estimation routine are also directed to the estimation recipe in Appendix C, which provides a detailed guide to the approach.

The first stage

To keep the exposition straightforward, we approximate the production function with a Cobb–Douglas technology.²⁴ We write

$$y_t = \beta_0 + \beta_k k_t + \beta_s l_t^s + \beta_u l_t^u + \beta_e e_t + \beta_f f_t + \beta_m m_t + \omega_t + \eta_t, \tag{10}$$

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 labour input, l_t^u is the log of the unskilled labour input, and m_t , f_t and e_t denote log-levels of materials, fuels and electricity.

We proceed as if materials were the proxy, rewriting (10) as

$$y_t = \beta_s l_t^s + \beta_u l_t^u + \beta_e e_t + \beta_f f_t + \phi_t(m_t, k_t) + \eta_t, \tag{11}$$

where

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

As with Olley–Pakes, (11) can be estimated using OLS (and a polynomial expansion in m_t and k_t to approximate $\phi_t(\cdot)$).

We take an alternative approach to explore a different non-parametric estimator. We first estimate the conditional moments $E(y_t | k_t, m_t)$, $E(l_t^u | k_t, m_t)$, $E(l_t^s | k_t, m_t)$, $E(e_t | k_t, m_t)$, and $E(f_t | k_t, m_t)$ by regressing y_t (for example) on k_t and m_t .²⁵ We use a locally weighted quadratic least squares approximation, although in principle one could use any consistent parametric or non-parametric estimator for each of these conditional means.²⁶ We then subtract the expectation of (11) conditional on (k_t, m_t) from (11) to obtain

$$y_t - E(y_t | k_t, m_t) = \beta_s (l_t^s - E(l_t^s | k_t, m_t)) + \beta_u (l_t^u - E(l_t^u | k_t, m_t)) + \beta_e (e_t - E(e_t | k_t, m_t)) + \beta_f (f_t - E(f_t | k_t, m_t)) + \eta_t. \tag{12}$$

No-intercept OLS is then used on this equation to estimate first-stage parameters.

This completes the first stage. Although there are several estimation steps in a more general non-parametric approach like ours, no single step is more complicated than a (locally weighted) least squares regression.²⁷ If we were only concerned with the marginal productivities of the variable inputs (except the coefficient on the proxy variable), we could stop here. To obtain the capital coefficient, a plant-level measure of productivity, and/or an estimate of returns to scale we need a more complete model for $\phi_t(\cdot)$ because both electricity and capital enter it twice.

24. As Olley–Pakes note, the approach applies to quite general production technologies.

25. See Section 5.2, and especially the first few pages of Section 5.2.1 of Pagan and Ullah (1999), for a (relatively) understandable discussion of kernel-based estimates of the coefficients on the linear terms of the estimating equation.

26. Readers not familiar with local quadratic 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, m_t) using as regressors the basis for a second-order polynomial approximation in (k_t, m_t) . For any particular point (k_t^*, m_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^*, m_t^*) most heavily. A consistent estimator of $E(y_t | k_t = k_t^*, m_t = m_t^*)$ is the intercept from this local quadratic regression.

27. We have used the OLS-with-a-polynomial-approximation approach (allowing for different sub-periods of the sample according to macroeconomic cycles) and we find in most cases that a third-order polynomial approximation gives very similar estimates of the parameters.

The second stage

We use two moment conditions to identify β_m and β_k . As with Olley–Pakes, our first moment condition identifies β_k by assuming that capital does not respond to the innovation in productivity ξ_t . The second moment identifies β_m by using the fact that last period's materials choice should be uncorrelated with the innovation in productivity this period. These population moments are given by

$$E[(\xi_t + \eta_t)k_t] = E[\xi_t k_t] = 0, \quad (13)$$

and

$$E[(\xi_t + \eta_t)m_{t-1}] = E[\xi_t m_{t-1}] = 0. \quad (14)$$

We obtain an estimate of the residual from the following relationship:

$$\xi_t + \hat{\eta}_t(\beta^*) = y_t - \hat{\beta}_s l_t^s - \hat{\beta}_u l_t^u - \hat{\beta}_e e_t - \hat{\beta}_f f_t - \beta_m^* m_t - \beta_k^* k_t - E[\omega_t | \omega_{t-1}],$$

where we explicitly reference the residual as a function of the two parameters $\beta^* = (\beta_m^*, \beta_k^*)$. To estimate $E[\omega_t | \omega_{t-1}]$ we use the estimates of ω_t obtained from the first stage results (and the candidate values (β_m^*, β_k^*)).²⁸

We also include the six over-identifying conditions, yielding in total eight population moment conditions given by the vector of expectations

$$E[(\xi_t + \eta_t)Z_t],$$

where Z_t is the vector given by $Z_t = \{k_t, m_{t-1}, l_{t-1}^s, l_{t-1}^u, e_{t-1}, f_{t-1}, k_{t-1}, m_{t-2}\}$. Finally, we obtain estimates $(\hat{\beta}_k, \hat{\beta}_m)$ by minimizing the GMM criterion function

$$Q(\beta^*) = \min_{\beta^*} \sum_{h=1}^8 \left(\sum_i \sum_{t=T_{i0}}^{T_{i1}} (\xi_{i,t} + \hat{\eta}_{i,t}(\beta^*)) Z_{i,ht} \right)^2, \quad (15)$$

where i indexing firms is explicit, h indexes the eight instruments, and T_{i0} and T_{i1} index the second and last period firm i is observed.

Inference using the bootstrap

Measuring the precision of our estimates requires us to account for the variance in every estimator that enters our routine (and all of their covariances). There are 11 estimating equations in total, and many estimates get used more than once. Pakes and Olley (1995) provide the theoretical framework for computing asymptotic standard errors.

Instead of undertaking this difficult task, we use the bootstrap for inference.²⁹ This technique (re)samples the empirical distribution of the observed data to construct new “bootstrapped” samples. The value of the statistic is computed for each of these samples, and the distribution of estimates so generated provides the bootstrap approximation to the true sampling distribution of the statistic.

Our resampling rule treats each set of firm-level observations together as an independent, identical draw from the overall population of firms. We sample with replacement and with equal probability from the sets of firm-level observations in the original sample. A bootstrap sample is considered complete when it has a number of firm-year observations that equals (or just exceeds) the number of firm-year observations in the original data.

The bootstrap is easy to implement. In addition, it also provides asymptotic refinements for many statistics, including the asymptotically pivotal ones in our analysis. Finally, the

28. See Appendix C.

29. See Horowitz (2001) for an overview of the bootstrap.

bootstrap makes inference on *differences* between estimators remarkably straightforward. The usual difficulty when constructing an estimate of the variance of differences is the need for the covariance term between the two estimators (they are estimated on the same sample). A distribution of differences obtains *across* the bootstrapped samples by subtracting one estimate from the other (for each of these samples). The sampling distribution so obtained automatically accounts for the covariance between the estimators.

The bootstrap approach must be slightly modified when using more moments than parameters to obtain estimates (as we do with the test of over-identifying conditions). The logic of the bootstrap requires that estimates obtained using bootstrapped samples must implement moments that equal zero in the population from which the bootstrap samples are drawn (that is, the observed data). Since this population is the original data, the bootstrapped moments have to be recentered by the estimated values of the moments using the original data (at the objective function minimum).

The Bond and Blundell IV approach

An alternative IV estimation strategy that also deals with the core issue of simultaneity is proposed by Blundell and Bond (2000). They start with the following model (in their notation):

$$y_{it} = \beta' x_{it} + \gamma_t + (\eta_i + v_{it} + m_{it}),$$

where y and x are (log) output and inputs (same as above), γ_t is a time-specific effect, η_i is a firm-specific fixed effect, v_{it} is AR(1), and m_{it} is MA(0) (say) arising from measurement error. In this model η_i , v_{it} and m_{it} can all potentially result in estimates that are biased. Their IV estimator is robust to correlation between each of these errors and potentially mismeasured input choices (at the expense of placing significant demands on the data).

They use two kinds of moments for identification. The first set uses input levels lagged at least two periods as instruments in the first-differenced equations (where output is also first-differenced to condition on the AR(1) productivity term). They report difficulty in obtaining precise estimates using just these moment conditions. They add an additional set of moments that uses “suitably lagged first differences of variables as instruments for the equations in levels”. The additional moments lower the standard errors and pass an over-identifying test.

We implement the Bond and Blundell estimator as an alternative to our approach and report the results in the next section.

6. RESULTS

In this section, we present several sets of results with several objectives in mind. Our over-riding goal is to illustrate how one can most usefully implement the intermediate inputs approach. No one approach will be appropriate for all industries in all timeframes. Instead, the *right* approach will depend on the details of the industry being studied. Towards this end, we show the reasoning used to select among proxies.

We find using either materials or electricity as a proxy yields statistically significant estimates of the parameters of production functions in the Chilean case. The estimates highlight how estimators using intermediate inputs to control for unobservables differ in predictable and informative ways from other existing and commonly used estimators. Our results are fairly robust across the industries we examine.

The base case

We begin by presenting production function estimates for the four industries discussed in Section 4—Food Products (311), Metals (381), Textiles (321) and Wood Products (331). Table 3

TABLE 3
Base case parameter estimates for four industries
(bootstrapped standard errors in parentheses)

Input	Industry (ISIC code)			
	311	381	321	331
Unskilled labour	0.139 (0.010)	0.172 (0.033)	0.130 (0.024)	0.193 (0.034)
Skilled labour	0.051 (0.009)	0.188 (0.025)	0.155 (0.026)	0.133 (0.030)
Electricity	0.085 (0.007)	0.081 (0.015)	0.005 (0.019)	0.047 (0.021)
Fuels	0.023 (0.004)	0.020 (0.011)	0.038 (0.010)	0.021 (0.014)
Materials	0.500 (0.078)	0.420 (0.091)	0.500 (0.118)	0.550 (0.086)
Capital	0.240 (0.053)	0.290 (0.094)	0.180 (0.095)	0.190 (0.090)
Returns to scale	1.037 (0.059)	1.172 (0.075)	1.007 (0.113)	1.133 (0.157)
No. obs.	6115	1394	1129	1032

presents the results using materials as the intermediate input proxy.³⁰ We find that coefficients are precisely estimated at standard levels of statistical significance. (The sole exception is the coefficient on electricity in ISIC 321.) Especially in the largest industry (ISIC 311), estimates are quite precise.

There are significant differences in the production functions across these four industries, but none of the industries really stands out as having a radically different technology. In all industries, the coefficient on materials is the largest and consistently hovers around 0.50. Capital is usually the factor with the next highest coefficient. Returns to scale range from 1.04 (for ISIC 311) to 1.172 (for ISIC 381), although estimates are generally not significantly different from constant returns. We will indirectly return to these base case results since many of our concerns are focused not on the coefficients in Table 3 *per se*, but rather how these coefficients differ from those obtained with traditional estimators.

Alternative proxies

The results in the base case use materials as the intermediate input proxy. There are, though, other candidate intermediate inputs, and these include fuels and electricity. In Section 4, we discussed why we prefer materials and electricity to fuels as our proxy. The main reason is both are non-zero for virtually all firms for all time periods (basically eliminating the need to truncate observations). Section 4 also suggested three other specification tests for the choice of the proxy.

30. See Levinsohn and Petrin (2000) for results from these four industries and four others with electricity as the proxy. Levinsohn and Petrin (1999) provides estimates of value-added production functions and implied productivity numbers also using the electricity proxy.

The results of each of these three tests are reported here. In each test, we take materials as the principle proxy and electricity as the other candidate.

The first specification test involves visually examining the function $\omega_t = \omega_t(m_t, k_t)$. Recall that the monotonicity condition requires that this function be increasing in materials (conditional on capital). The first specification test simply graphs the smoothed ω -function and looks for this monotonicity. Because we believe that this function may differ over the three macroeconomic cycles in our data, we have three such functions to graph. The first spans the years 1979–1981, the second 1982–1983, and the third 1984–1986.

A good example of the results that come out of this exercise is provided by ISIC 321 (textiles). The three panels of Figure 1 show the smoothed plots for materials in this industry. The vertical axis measures the estimated productivity shock, while the axis running left measures materials usage and the axis running right measures capital. If the monotonicity condition always held, conditional on any observed level of capital materials usage would increase when productivity increased. In the first panel, productivity is indeed increasing in materials for all levels of capital. This also appears to be the case in the second time period. In the third period, at low levels of materials usage and high levels of capital, the monotonicity condition is sometimes violated. Overall, we find that monotonicity appears to largely hold for the biggest three industries for all three periods.³¹

In any particular industry these functions often differ across the periods (as in ISIC 321). Additionally, within an industry-time period, the rates of increase for productivity appear to vary widely across capital levels (also as in ISIC 321). These results imply that the non-parametric approaches are important; they provide a flexibility robust to these differences. Of course, nothing in our methodology precludes these non-parametric plots from looking like the Andes—full of (smoothed) peaks and valleys—and we find this to be the case for our smallest industry for the last two time periods. For this industry one could undertake some of the specification changes suggested in Section 4 to see if they restore monotonicity.

Our second specification test compares the parameter estimates obtained with different proxies. We simply ask if the choice of proxy matters for the resulting estimates. If the model is correct, any proxy satisfying the monotonicity condition should theoretically give similar parameter estimates for the freely variable inputs (other than the intermediate proxy). Put another way, failure of the first stage estimates to be invariant to the choice of the proxy means either one (or both) of the proxies are invalid, or the overall model is incorrect.

The top panel of Table 4 lists the differences in estimates between electricity and materials (electricity minus materials). We find that electricity yields almost identical estimates. A 90% symmetric confidence interval constructed directly from the bootstrapped distribution reveals that none of the differences are statistically significant at 10%.³² By this measure, the choice between electricity and materials does not appear to make a difference.

The third specification test uses over-identifying restrictions to explore the consistency of the model and data. As described earlier, we ask whether inputs lagged one period are correlated with this period's innovation in productivity (again, we must lag the intermediate input and capital two periods). If the proxy is conditioning out all of the variation in inputs that is correlated with the productivity shock, the value of the objective function obtained using the actual sample should not differ markedly from the values obtained using the (recentred) bootstrap samples (*i.e.* the over-identifying conditions should hold).

31. We also find this to be true when using electricity.

32. See Hall (1986) for encouraging results on the number of bootstrap replications necessary to obtain reasonable coverage probabilities of confidence intervals.

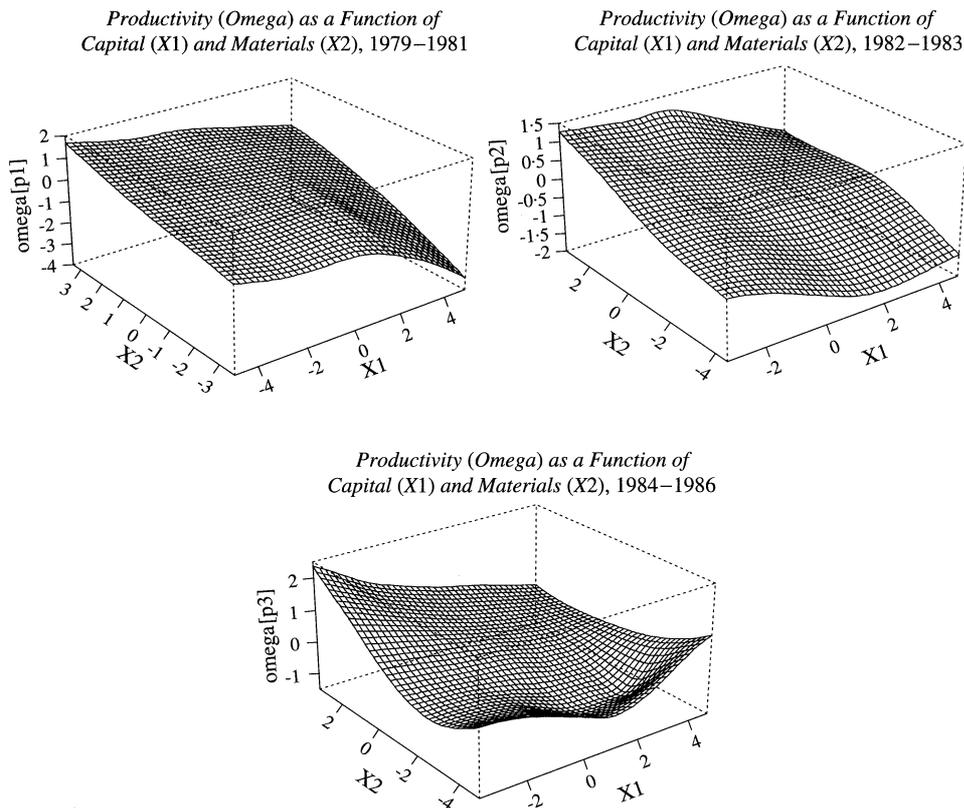


FIGURE 1

Productivity as a function of capital and materials use

The bottom row of Table 4 reports the P -values for a test of the null hypothesis that the over-identifying restrictions hold for each proxy in each industry. Here we find that for three of the four industries we do not reject at reasonable levels. However, we do reject in our largest industry (ISIC 311) at any significance level greater than 1% for both materials and electricity. In part this result may obtain because of the additional power achieved by the fivefold increase in observations (relative to the smaller industries). Loosening the functional form restrictions may be helpful in getting the over-identifying conditions to hold here. In particular, one might add higher-order terms of inputs and/or interaction terms between inputs to the first stage of the estimation, especially if the rejection can be tied directly to a particular input (or inputs). Additionally, one might try loosening the restriction on the ϕ function, allowing it to vary by other observables (like time, firm type or location).

Alternative estimators

In this section, we compare our base case results with those obtained using alternative estimators. We start with OLS, the simplest and most often used alternative. We are especially interested in investigating whether our results are consistent with the story of a proxy controlling for an unobserved transmitted productivity shock. Marschak and Andrews suggested that the transmitted productivity shock would be positively correlated with variable inputs. In the OLS case, esti-

TABLE 4
Two specification tests for the proxy choice electricity and materials

Difference in point estimates	Industry (ISIC code)			
	311	381	321	331
Unskilled labour	-0.002 (0.005)	-0.004 (0.009)	-0.001 (0.012)	0.010 (0.014)
Skilled labour	-0.003 (0.003)	-0.004 (0.008)	-0.011 (0.012)	0.006 (0.011)
Fuels	0.000 (0.002)	0.004 (0.004)	0.002 (0.004)	-0.005 (0.006)
Capital	0.03 (0.075)	-0.03 (0.131)	0.07 (0.145)	0.13 (0.109)
Number sig. different at 10%	0	0	0	0
<i>P</i> -value for over-identification test				
Materials	0.01	0.51	0.86	0.68
Electricity	0.00	0.32	0.10	0.40

Note: The first panel lists the difference in the coefficients on the non-proxy inputs using electricity vs. materials as the proxy (electricity estimates minus materials estimates). Bootstrap standard errors of the differences are reported in parentheses. The number of these differences that are statistically different from zero at a level of significance of 0.1 is computed using a symmetric 90% confidence interval implied by the bootstrap. The second panel lists the *P*-value for the test of the over-identifying restrictions for each proxy in each industry.

mates of the coefficients on the variable inputs are likely to be biased upward (see Section 2). 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, the OLS estimate on capital is likely to be biased downward.

Figure 2 provides six histograms pertaining to ISIC 311 (the largest industry). We initially focus our attention on the first histogram (top left corner). This graph gives the empirical distribution of the difference between the OLS estimate and our (LP) estimate of the coefficient on unskilled labour, β_u . The histogram is constructed in the following manner. The production function is estimated 200 times using LP (each time using a new bootstrapped sample of firms and recentred moments). With each of these same 200 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 on average there were no differences between the OLS and LP estimates, we would expect to see the histograms for the variable factors centred symmetrically around zero (with approximately 50% of the samples to the right and the left of zero). For this particular coefficient, the OLS estimate is larger than the LP estimate 70% of the time.

OLS and LP starkly differ on the estimate of the skilled labour coefficient (top right corner). For every one of the 200 samples, the coefficient on skilled labour was larger when OLS was used compared to the LP estimates.³³ The results are similarly stark for the histograms of estimates

33. The size of the difference is small—about 0.01—although the estimated coefficient is only 0.05 so the difference in percentage terms is about 20%.

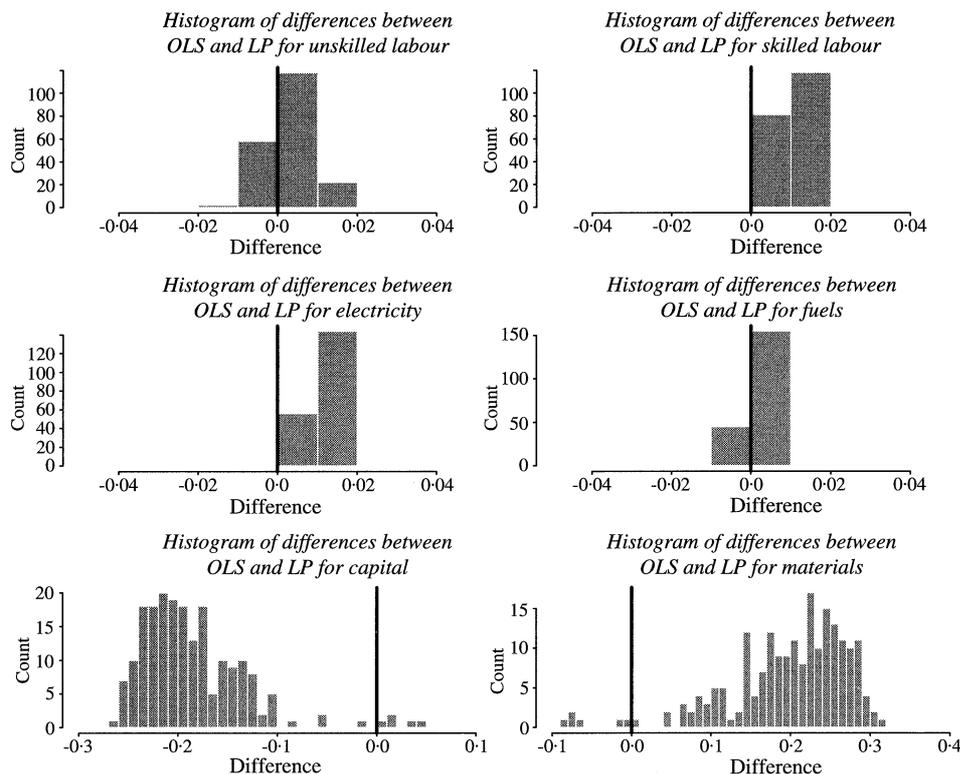


FIGURE 2
Differences between OLS and LP estimates

for electricity and materials, and they are slightly less so for fuels. Overall, for all of the variable inputs, the coefficients obtained with OLS are on average larger than those obtained with the intermediate input proxy estimator.

At the bottom left corner of Figure 2, we present the histogram of differences between the OLS and LP estimates for the capital coefficient. This histogram shows that for almost every one of the 200 times we estimated the production function, the OLS estimate of β_k was smaller than the LP estimate. Taken together with the histograms for the variable factors, the evidence seems consistent with Marschak and Andrews' concern. It also suggests that our input proxy is doing what it is intended to do, controlling for an unobservable transmitted shock that is highly correlated with freely variable inputs.

The pattern of biases in Figure 2 is the same as that found by Olley and Pakes in the single industry they examined (where they use investment as their proxy in a value-added production function). We wish to examine whether this pattern holds consistently across all four industries. To summarize the 24 histograms that this exercise yields, we report two summary statistics for each histogram in Table 5. The top row for each input gives the estimated difference ($\hat{\beta}_{OLS} - \hat{\beta}_{LP}$) between the OLS and LP estimates of that coefficient. The second row reports the per cent of the 200 samples where realized differences were greater than zero.

The pattern observed in ISIC 311 appears to carry over in spirit to the smaller industries. For all four industries together, 15 of the 20 coefficients estimated on the variable inputs are greater for OLS vs. LP, both in terms of the difference in the point estimate (on the original sample), and

TABLE 5
OLS estimate minus the LP estimate

Input	Industry (ISIC code)			
	311	381	321	331
Unskilled labour				
Difference	0.004	0.001	-0.010	-0.009
% > 0	70	51	21	27
Skilled labour				
Difference	0.011	-0.008	-0.003	0.013
% > 0	100	23	48	84.5
Electricity				
Difference	0.011	0.005	0.005	0.007
% > 0	100	74.5	83.5	68.5
Fuels				
Difference	0.001	0.004	0.002	-0.004
% > 0	77.5	81	61	20.5
Materials				
Difference	0.210	0.166	0.197	0.075
% > 0	97	76.5	84.5	88.5
Capital				
Difference	-0.190	-0.225	-0.141	-0.112
% > 0	2.5	14.5	20	24
Returns to scale				
Difference	0.049	-0.056	0.049	-0.030
% > 0	78.0	25.0	64.5	56.5
No. obs.	6115	1394	1129	1032

Note: The top entry in each cell is the difference between the OLS and LP estimate. The bottom entry in each cell is the percentage of the 200 bootstrapped samples that yielded estimates in which the OLS coefficient was larger than the LP coefficient (*i.e.* $\hat{\beta}_{OLS} - \hat{\beta}_{LP} > 0$).

in the per cent of differences greater than zero obtained from the bootstrap. Electricity, materials, and fuels provide the strongest evidence of being overestimated. The OLS capital estimates, on the other hand, all fall below their equivalent LP estimate. Overall, 19 of the 24 differences in coefficients are consistent with the simultaneity story (11/12 if we focus on the two largest industries). Thus, 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 differences varies by factor and by industry, but the flavour of the results tend to be most convincing in the industries with the smallest sampling error.

The bottom of the table presents 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 not obvious whether OLS will over- or underestimate returns to scale. Two industries have point estimates for returns to scale that are larger for OLS than LP, and two industries have point estimates for which the opposite is true. From the bootstrapped samples it appears that in three of four industries OLS is more likely to produce larger estimates of returns to scale than LP.

We also investigate whether the Olley–Pakes (OP) estimates are the same as the LP estimates. Figure 3 reports the six histograms for ISIC 311 which compare these two estimators. The picture is fairly similar to that obtained under OLS, with the exception of the unskilled labour

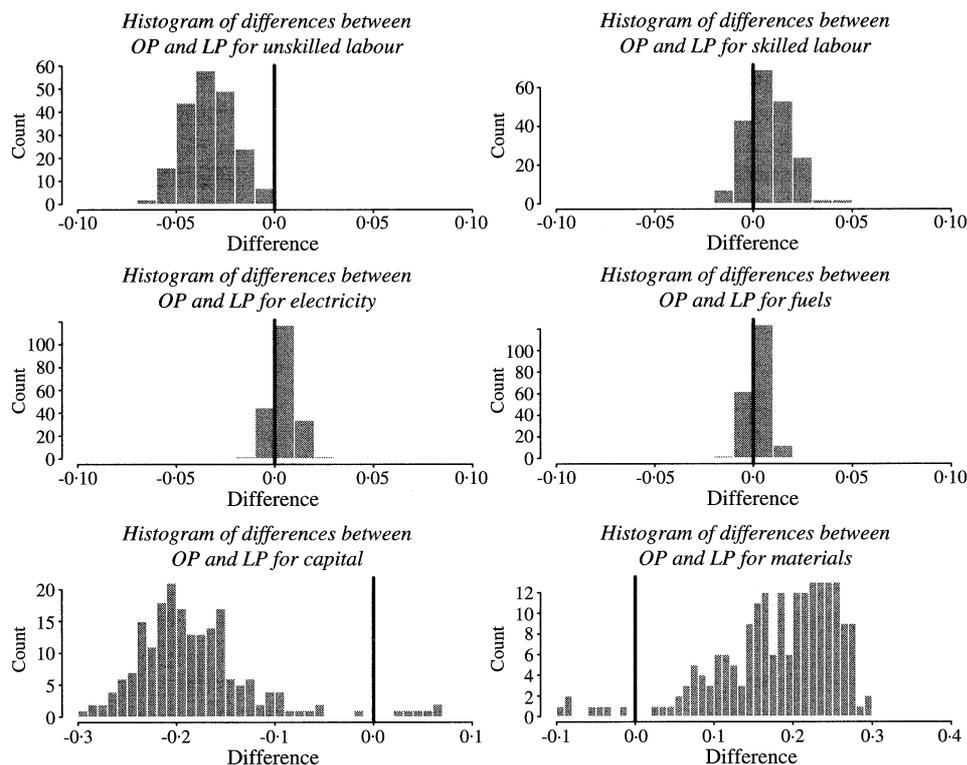


FIGURE 3

Differences between OP and LP estimates

coefficient. Thus, Figure 3 is consistent with the intermediate input responding more fully to the productivity shock than investment.

OLS and Olley–Pakes are but two alternatives to our approach. Other commonly used alternatives include fixed effects and instrumental variables (where inputs lagged one period are used as instruments). We compare estimates across all of these alternatives in Table 6. In order to get some sense of how truncating on positive investment affects the estimates, we also compute LP on just firms with positive investment. Finally, we estimate parameters using the Blundell and Bond (2000) approach.

Before proceeding to the table we note that we do not report results for Blundell and Bond. The parameter estimates we obtained using this estimator were very sensitive to the particular bootstrapped sample upon which we were estimating. This led to many estimates obtaining at the boundary of parameter space (which we took to be the unit interval for each of the six parameters). Since the sampling theory for estimators at the boundary is complicated (and generally not normal), we do not compare this estimator to the others. We simply note that very little variation remained in our data after all of the possible estimation problems under Blundell and Bond were accounted for (and all of the associated variation was conditioned out).

Since the estimators are generally not nested, we simply report in Table 6 results from tests for significant differences between estimates. For any comparison, the table contains the *P*-value of the “no differences” null hypothesis (where the statistic is computed by pre- and post-multiplying the variance–covariance matrix of the estimated differences by the estimated differences).

TABLE 6
 Comparisons across estimators P -value for $H_0: \beta_1 = \beta_2$

Comparison	Industry (ISIC code)			
	311	381	321	331
Levinsohn–Petrin vs.				
OLS	<0.01	0.20	0.58	0.21
Fixed effects	<0.01	<0.01	<0.01	<0.01
Instrumental variables	<0.01	0.22	0.09	<0.01
Olley–Pakes	<0.01	0.54	0.20	0.89
Levinsohn–Petrin ($i > 0$ only)	<0.01	0.02	0.27	0.93
Olley–Pakes vs.				
OLS	<0.01	0.04	0.19	0.46
Fixed effects	<0.01	<0.01	<0.01	<0.01
Instrumental variables	<0.01	<0.01	<0.01	<0.01
Levinsohn–Petrin ($i > 0$ only)	0.56	0.47	0.85	0.55
Fixed effects vs.				
OLS	<0.01	<0.01	<0.01	<0.01
Instrumental variables	<0.01	<0.01	<0.01	<0.01
No. obs.	6115	1394	1129	1032

Note: The cells in the table contain the P -value for a standard Wald test for “no differences between the (vector of) parameter estimates for estimators 1 and 2”. <0.01 indicates a P -value that is less than 0.01.

We start with industry ISIC 311. This industry, which is the largest industry, yields the most disagreement between estimators. In this industry, LP is not in agreement with OLS, IV, fixed effects, LP run on firms with positive investment, or OP; they all reject the null of compatibility at a 1% level of significance. Similar results obtain when OP is compared to OLS, IV, and fixed effects, and when fixed effects is compared to OLS and IV. The *only* pair of estimators that do not disagree is OP and LP estimated on firms with positive investment.

Comparing estimates across industries, it appears that LP strongly disagrees with IV, rejecting at 10% in three of the four industries, and suggesting last period’s inputs levels respond to the same unobserved to which this period’s inputs also respond. LP also disagrees with LP estimated on firms with positive investment in the two biggest industries (rejecting at 5%). We think this suggests that truncating observations with zero investment may not generally be innocuous. However, we do find that LP and OP only strongly disagree in ISIC 311, suggesting the persistence that they each correct for has a reasonable common component.

Moving down the table to comparisons with OP, we find this estimator strongly disagrees with fixed effects and IV, rejecting at 1% in all four industries. OP also stands in contrast to OLS, differing significantly at 5% in the two biggest industries. OP and LP using just positive investment appear to be the most compatible with each other, again suggesting a commonality between these estimators.

We conclude our discussion by noting that the fixed effects estimator is in the most pronounced disagreement with the other estimators as it rejects compatibility with *every* other estimator in every industry at the 1% level of significance. Previous results suggest the existence of a persistent shock which is highly correlated with input choices. This result says that this persistent shock seems to vary within-firm over time.

7. CONCLUSIONS

In this paper we add to the methods that control for correlation between input levels and an unobserved firm-specific “productivity” process. Olley and Pakes show how to use investment

to do so, and we extend their idea to other inputs. We discuss some theoretical benefits of extending the proxy choice set in this direction. Our empirical results suggest these benefits can be important.

There is not a universally best way to implement the proxy approach. The choice of the proxy, be it investment or an intermediate input, depends on the details of the data and the industry. It is important, for example, that the selected proxy be reported and non-zero for most firms in most years. It is also important that the data substantiate, to a large extent, the monotonicity condition. In this paper, we describe these and other methods which can provide guidance when evaluating the appropriateness of any particular proxy. To this effect our emphasis has been less on particular results and more on how to implement the proxy approach.

We report several different findings. We usually reject that the intermediate input proxy gives the same results as the more traditional OLS, IV, or fixed effects estimators. The ways in which these traditional estimators and our estimator differ are predictable and consistent with the economics underlying our approach. We also find that the estimates from our approach and that of Olley and Pakes also differ, but the difference is not as dramatic as that between our estimator and the more traditional estimators. We do not find this surprising, given that the Olley and Pakes approach is a very useful one for addressing simultaneity problems. Our approach is a more subtle change in implementation vis-à-vis Olley and Pakes. Its usefulness derives from the fact that it provides an alternative that is easy to implement, it allows the researcher to use more of the existing data, it generally works well in practice, and it appears to address some situations in which the OP approach may not work well.

This paper has not explored any of the implications of the production function estimates. For example, 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.

Here our objective has been to establish a framework for using proxies to address simultaneity problems. There are many other remaining issues that are beyond the scope of our approach here. We hope these unanswered questions will not deter researchers from implementing what we believe is a pretty simple (as simple as OLS, if you like) and straightforward approach to addressing the simultaneity problem. In the end, the importance of correcting for simultaneity will be determined by the particular application one has in mind. Most of the evidence in this paper suggests that when it comes to estimating production functions, addressing Marschak and Andrews' simultaneity concern is important.

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 $\iota(\omega; p_L, p_I, K)$ that is strictly increasing in productivity (ω) (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 labour (L) and the intermediate input (ι), and $f_{L\omega}$, $f_{\omega\omega}$, 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 labour 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 labour 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 ω , and when more than one type of labour exists.

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

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

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

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

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

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

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

so

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

APPENDIX B

A short-cut

If the production function is weakly separable in an intermediate input that satisfies the monotonicity condition, and cost-minimization and perfect competition hold, the contribution of the separable input to output is given by its revenue share (s_t) times its level, or $\beta_t \iota_t = s_t \iota_t$ (see for example Solow (1957) or Griliches and Ringstad (1971)). Use of this restriction leaves only β_k to be estimated in the second stage since $s_t \iota_t = \beta_t m_t$ is net out of y_t in the first stage. (8) becomes

$$y_t^* = \beta_k k_t + E[\omega_t | \omega_{t-1}] + \eta_t^*,$$

where $\eta_t^* = \xi_t + \eta_t$ and $y_t^* = y_t - l_t \beta_l - s_t \iota_t$.³⁴ Now the second stage has only one parameter to be estimated, β_k . This restriction can significantly reduce the computational burden and improve efficiency (if it holds).

Does this short-cut make good sense in practice? The answer depends on how comfortable one is with the separability assumption. In practice, whenever one estimates a value-added production function, separability is assumed for all the intermediate inputs that are subtracted out.³⁵ This short-cut only requires separability in the one intermediate input used as the proxy, and so is less restrictive than what is assumed with the typical value-added production function.

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. The recipe below is not written with any particular software package in mind, and we have used (on their own) Stata, Gauss, and SPlus to implement variants of this routine.

34. The intercept is suppressed.

35. See Bruno (1978) and Basu and Fernald (1995) on estimation using value-added production functions.

ESTIMATION RECIPE

Stage one:

1. Run a regression of y_t on m_t and k_t to obtain an estimate of the function $E(y_t | m_t, k_t)$ (we use a locally weighted least squares regression).
2. Run a regression of l_t^u on m_t and k_t to obtain an estimate of the function $E(l_t^u | m_t, k_t)$.
3. Repeat step 2 to obtain estimates of the functions $E(l_t^s | m_t, k_t)$, $E(e_t | m_t, k_t)$, and $E(f_t | m_t, k_t)$.
4. Construct $Y(e_t, k_t) = y_t - E(y_t | m_t, k_t)$ using the estimate of the conditional expectation from the regression in step 1. This is the dependent variable in step 5. Similarly, difference out the predicted mean for each of the explanatory variables, and call these new regressors that are net of materials and capital variation ($X_1(m_t, k_t)$, $X_2(m_t, k_t)$, $X_3(m_t, k_t)$, $X_4(m_t, k_t)$).
5. 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. The key estimated parameters from this stage are the production function parameters on all the variable inputs except the intermediate proxy, or $\hat{\beta}_u$, $\hat{\beta}_s$, $\hat{\beta}_e$, and $\hat{\beta}_f$.

Stage two:

1. Compute the estimate of $\phi_t(m_t, k_t)$ for each of the three different time periods 1979–1981, 1982–1983, and 1984–1986. To do so use the appropriate observations and (some form of) regression to predict $y_t - \hat{\beta}_s l_t^s - \hat{\beta}_u l_t^u - \hat{\beta}_e e_t - \hat{\beta}_f f_t = \phi_t + \eta_t$ using (m_t, k_t) as explanatory variables. Save the estimate $\phi_t(\cdot)$.
2. Choose a candidate value for (β_m, β_k) , say (β_m^*, β_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 $\omega_t + \eta_t = y_t - \hat{\beta}_u l_t^u - \hat{\beta}_s l_t^s - \hat{\beta}_e e_t - \hat{\beta}_f f_t - \beta_m^* m_t - \beta_k^* k_t$. For notation’s sake, call the variable just computed “A”.
4. Compute $\omega_{t-1} = \phi_{t-1} - \beta_k^* k_{t-1} - \beta_m^* m_{t-1}$. Call this variable “B”.
5. Regress “A” on “B” (again we use locally weighted least squares). Call the predicted values “C”. “C” is an estimate of $E(\omega_t | \omega_{t-1})$.
6. Compute $(\xi_t + \eta_t)$ by substituting “C” in for $E[\omega_t | \omega_{t-1}]$ to obtain

$$\xi_t + \eta_t(\beta_m^*, \beta_k^*) = y_t - \hat{\beta}_s l_t^s - \hat{\beta}_u l_t^u - \hat{\beta}_e e_t - \hat{\beta}_f f_t - \beta_m^* m_t - \beta_k^* k_t - E[\omega_t | \omega_{t-1}].$$

This is the residual that enters the moment equation. Use it to construct the sample analogues to the population moment conditions.

7. Using your favourite minimization routine, choose $(\hat{\beta}_m, \hat{\beta}_k)$ to minimize the objective function from (15) (i.e. distance between the observed moments and zero). This will entail iterations over the previous six steps.

Acknowledgements. We would like to thank our referees and Editor for helpful suggestions that led to a substantially revised paper. We would like to thank seminar participants at UC Berkeley, University of Toronto, Yale University, Harvard University, University of Chicago, and NBER for helpful suggestions on earlier work on this project. Jason Abrevaya, Dan Akerberg, Susanto Basu, Joel Horowitz, Peter Klenow, Steve Olley, Ariel Pakes and Mark Roberts 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.

REFERENCES

- ATTANASIO, O., PACELLI, L. and DOS REIS, I. R. (2000), “Aggregate Implications of Plant Level Investment Behavior: Evidence from the U.K. and.” (Working Paper, University College London).
- BASU, S. and FERNALD, J. (1995), “Are Apparent Productive Spillovers a Figment of Specification Error?”, *Journal of Monetary Economics*, **36** (1), 165–188.
- BITRAN, E. and SAEZ, R. (1994), “Privatization and Regulation in Chile”, in B. Bosworth, R. Dornbusch and R. Laban (eds.) *The Chilean Economy: Policy Lessons and Challenges* (Washington, D.C.: The Brookings Institution) 329–377.
- BLUNDELL, R. and BOND, S. (2000), “GMM Estimation with Persistent Panel Data: An Application to Production Functions”, *Econometric Reviews*, **19** (3), 321–340.
- BRUNO, M. (1978) *An Analysis of the Concept of Real Value-Added, in Production Economics: A Dual Approach to Theory and Applications*, Chapter III.1 (North-Holland).
- CHAMBERS, R. G. (1997) *Applied Production Analysis* (Cambridge University Press).

- DOMS, M. and DUNNE, T. (1998), "Capital Adjustment Patterns in Manufacturing Plants", *Review of Economic Dynamics*, **1**, 409–429.
- 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.
- FLUX, A. W. (1913), "Gleanings from the Census of Production Report", *Journal of the Royal Statistical Society*, **76** (6), 557–598.
- GRILICHES, Z. and MAIRESSE, J. (1998), "Production Functions: The Search for Identification", in *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium* (Cambridge University Press) 169–203.
- GRILICHES, Z. and RINGSTAD, V. (1971) *Economies of Scale and the Form of the Production Function* (North-Holland).
- HALL, P. (1986), "On the Number of Bootstrap Simulations Required to Construct a Confidence Interval", *Annals of Statistics*, **14** (4), 1453–1462.
- 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).
- HOROWITZ, J. (2001), "The Bootstrap", in J. J. Heckman and E. Leamer (eds.) *Handbook of Econometrics*, Vol. 5, (Oxford: Elsevier Science) 3159–3228.
- LEVINSOHN, J. (1999), "Employment Responses to International Liberalization in Chile", *Journal of International Economics*, **47**, 321–344.
- LEVINSOHN, J. and PETRIN, A. (1999), "When Industries Become More Productive, Do Firms? Investigating Productivity Dynamics" (NBER Working Paper 6893).
- LEVINSOHN, J. and PETRIN, A. (2000), "Estimating Production Functions Using Inputs to Control for Unobservables", (NBER Working Paper 7819).
- 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 M. Roberts and J. Tybout (eds.) *Producer Heterogeneity and Performance in the Semi-Industrialized Countries*, Chapter 4 (World Bank).
- MARSCHAK, J. and ANDREWS, W. H. (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 C. Sims (ed.) *Advances in Econometrics, Sixth World Congress*, Vol. II (Cambridge University Press) 171–259.
- 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), 545–564.
- 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.
- VARIAN, H. R. (1992) *Microeconomic Analysis* (W. W. Norton).