Identifying Production Functions Using Restrictions from Economic Theory

University of Wisconsin-Madison

November 10, 2008

1 Introduction

As first pointed out by Marschak and Andrews [1944], using the inputs and outputs of profit maximizing firms to estimate production functions gives rise to an endogeneity problem. The endogeneity problem is caused by the presence of productive factors that are unobservable to the econometrician but that are "transmitted" to the firm's optimal choice of inputs. These unobservable factors are traditionally captured by a scalar productivity index that varies across firms and potentially evolves over time. The two traditional and oldest methods of controlling for the endogeneity problem in the production function are instrumental variable and fixed effect estimation. However these solutions have proven unsatisfactory on both theoretical and empirical grounds (for a review, see Griliches and Mairesse [1998] and Ackerberg et al. [2007]).

Instead, the modern literature on estimating production functions initiated by Olley and Pakes [1996] (OP for short) and further developed by Levinsohn and Petrin [2003] (LP for short) uses restrictions from economic theory to identify production function parameters. In particular, by assuming that productivity is the only dimension of unobserved heterogeneity among firms, these authors prove that inputs from the firm's profit maximization problem can be used to proxy for a firm's level of productivity and thereby solve the endogeneity problem. Following the language of Ackerberg et al. [2007], we shall generically refer to this method of solving the endogeneity problem as the "proxy variable" approach, and we shall refer to the underlying assumption that productivity is the only dimension of unobserved heterogeneity as the "scalar unobservability assumption".

In this paper we show how to use the full information contained in the firm's profit maximization to identify production function parameters. The key to our approach is to recognize that there exists more information in the firm's profit maximization problem than is utilized by the proxy variable approach. By tapping into this information, we show how to build on the OP and LP estimators so as to solve the endogeneity problem in an econometrically simple way that does not require the scalar unobservability assumption. Since "massive" heterogeneity is a persistent empirical finding from firm level studies (see e.g., Bartelsman and Doms [2000]), we feel that being able to identify production function parameters in the presence of possibly complicated patterns of unobserved heterogeneity is an important contribution with many possible applications, a few of which we will sketch. Moreover as our identifying assumptions are fully consistent with the underlying economic assumptions implicit in the proxy variable approach, our estimation strategy does not contradict the OP and LP methodologies (which are now widely used by the applied community), but rather extends then by fully exploiting the economic restrictions in their models.

The plan of the proposal is as follows. We first introduce the underlying economic model used by OP and LP and review the scalar unobservability assumption. After examining the full economic implications of scalar unobservability, we then present our solution to the endogeneity problem, which is followed by an application to the Chilean plant level data. We then discuss a non-parametric testable implication of the Cobb-Douglas specification that we use and implement it against the Chilean data. Our findings suggest that functional forms beyond Cobb-Douglas, and in particular CES, may be more appropriate for the data. We also extend our methodology to handle the problem of so called "revenue production functions" created as a result of output being measured in such a way that it is confounded by prices. Finally we close with an outlook for further development of our theory and applications.

2 The Model

The first step in estimating production functions against plant level data is to write down the functional form of the production technology. We explicitly mention this seemingly trivial point for an important reason. While one of key strengths of the semiparametric approaches of OP and LP is that they avoid having to specify a number of the functional relationships that characterize the economic environment, the production function technology is not one of them. The key to our approach is to recognize that the actual form of the production function combined with calculus of the firm's profit maximization problem provides useful information that is sufficient to identify the production function parameters. Moreover this information is available without departing from the underlying economic assumptions of the OP and LP models.

We will illustrate both the logic of the proxy variable approach and our proposed extensions in the context of the most widely applied functional form, namely a Cobb-Douglas technology. However our identification argument readily generalizes to other functional forms including the CES and generalized Diewert. In a separate section, we present the analogue for the CES of a few of our main results for the Cobb-Douglas case.

Consider a firm j that in period t produces an output Y using capital K and labor L and intermediate inputs M. The firm's technology is given by

$$y_{jt} = \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + a_{jt} \tag{1}$$

where small letters denote logs. All differences in technology between firms are thus captured by the residual a_{jt}. For simplicity we refer to intermediate inputs M as though it was a single input, when in reality it consists of several possible inputs, most notably energy and materials.

Following Griliches and Mairesse [1998], the residual a_{jt} can be decomposed into three components, i.e.,

$$a_{jt} = \omega_{jt} + \epsilon_{jt} + \eta_{jt}. \tag{2}$$

In (2), ω_{jt} is a the firm's anticipated productivity that is part of j's information set $\mathfrak{I}_{j,t}$ in period t. On the other hand, ϵ_{jt} is a mean zero unanticipated productivity shock that is independent of $\mathfrak{I}_{j,t}$ and only realized by the firm at the end of the period. Finally, η_{jt} is a mean zero measurement error (brought about, for example, by deflating revenue in order to measure output), which does not enter into the firm's problem but is only part of the econometrician's problem. The component ω_{jt} is thus the traditional measure of total factor productivity as it captures the systematic technology differences across firms. Moreover ω_{jt} is the source of the endogeneity problem: the fact that ω_{jt} is part of $\mathfrak{I}_{j,t}$ will generally induce a correlation between the residual a_{jt} and the firm's inputs (l_{jt}, k_{jt}, m_{jt}) that prevents OLS from consistently estimating the parameter vector $\theta = (\alpha, \beta, \gamma)$ using (1).

The proxy variable approach solves the endogeneity problem by explicitly modeling the input demand behavior of the firms, which provides the econometrician with a mechanism through which productivity is transmitted to the inputs. A critical feature of this model concerns the timing of input decisions, which we now discuss (for a more in depth discussion, see Ackerberg et al. [2006]).

In period t, firm j takes its productivity ω_{jt} and capital stock K_{jt} to be state variables that are fixed for the period. Productivity evolves according to an exogenous Markov process $\Pr\left(\omega_{jt} | \{\omega_{j\tau}\}_{\tau=1}^{t-1}\right)$, and capital is accumulated each period based on last period's investment decision $I_{j,t-1}$, and thus $K_{jt} = \mathbf{K}(K_{jt-1}, I_{jt-1})$. Intermediate inputs M_{jt} on the other hand are static and variable inputs. They are variable in the sense that can be adjusted each period, and they are static in the sense that they have no dynamic implications, i.e., M_{jt} does not affect the firm's profit in any future periods.

Labor is a less clear cut input. The usual assumption used in OP and LP is that L_{jt} is also a static and variable input, and thus treated exactly like the intermediate input from the firm point of view. However an alternative perspective is that even though labor may be variable each period, the presence of hiring/firing costs would cause labor to have dynamic implications, and thus lagged labor $L_{j,t-1}$ would enter as a state variable to the firm in period t. Alternatively, labor could be modeled with a "time to build" assumption similar to capital, due for example to the hiring/firing decisions in each period taking one or more periods to have effect. This would cause L_{jt} to be a fixed input in period t, and thus an element of the state space. In what immediately follows, we will pursue the interpretation actually employed by OP and LP, which is that labor is a variable and static input, and will address alternative interpretations of labor in the empirical application.

Given the state variables k_{jt} and ω_{jt}, the firm is assumed to have a profit function in period t of the form

$$\pi(k_{jt}, \omega_{jt}, \Delta_t) = \max_{l_{jt}, m_{jt}} \pi(l_{jt}, m_{jt}, k_{jt}, \omega_{jt}, \Delta_t), \tag{3}$$

where Δ_t is a state variable that summarizes the firm's economic environment. A critical feature of Δ_t is that it does not contain a firm subscript, and thus the economic environment is the same for all firms within a period. This is the scalar unobservability assumption, and it implies that all of the unobserved heterogeneity among firms is explained by heterogeneity in TFP.

By further assuming that the economic environment Δ_t and productivity ω_{jt} evolve in a first order Markov fashion, OP augment the "inner" maximization problem in (3) with an "outer" maximization problem over the firm's dynamic control, which is expressed through the Bellman equation[1]

$$V(k_{jt},\omega_{jt},\Delta_t) = \max_{i_{jt}\geq 0} \pi(k_{jt},\omega_{jt},\Delta_t) - c(i_{jt},\Delta_t) + \beta E[V(k_{jt+1},\omega_{jt+1},\Delta_{t+1}) \mid k_{jt},\omega_{jt},\Delta_t,i_{jt}], \quad (4)$$

where $c(i_{jt},\Delta_t)$ is the cost in period t of investing, and β is the discount factor.

Finally, assuming certain regularity conditions on the Bellman problem (4), Pakes [1994] shows that the firm's investment policy $i_{jt} = i(k_{jt},\omega_{jt},\Delta_t) = i_t(k_{jt},\omega_{jt})$ is a strictly increasing function of ω_{jt} over the range of strictly positive investment. Thus so long as $i_{jt} > 0$, the investment demand equation can be inverted to yield $\omega_{jt} = f_t(k_{jt}, i_{jt})$.

3 The OP Approach

The basic empirical strategy behind OP is to plug the inversion mapping $\omega_{jt} = f_t(k_{jt}, i_{jt})$ implied by their theory back into the production function to control for the endogeneity problem. Of course, to obtain an analytically closed form solution for the investment demand equation $i_t(k_{jt},\omega_{jt})$, or to even compute its value for any specification of the primitives, i.e., $\pi()$, $c()$, $Pr(\omega_{jt} \mid \omega_{jt-1})$, etc., would require a host of further functional form assumptions. Even then, the functional form of the investment demand equation would remain difficult to pin down because of the potentially complicated solution to the the dynamic programming problem. An important contribution of OP was to instead use the more basic restriction that there exists an inverse mapping from which a firm's productivity can be expressed as a function of its capital stock and its investment (so long as its investment is positive), and treat this function $f_t(k_{jt}, i_{jt})$ as a nuisance parameter to estimated nonparametrically along with the production function.

In particular, for all observations (j, t) for which $i_{jt} > 0$, we can substitute $f_t(k_{jt}, i_{jt})$ into

[1] We exclude the firm's entry/exit decision in each period as the main concern of the recent literature has been with the simultaneity problem rather than the selection problem.

the production function (1) to yield

$$\begin{aligned} y_{jt} &= \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + f_t\left(k_{jt}, i_{jt}\right) + \varepsilon_{jt} \\ &= \alpha l_{jt} + \gamma m_{jt} + \Phi_t\left(k_{jt}, i_{jt}\right) + \varepsilon_{jt} \end{aligned} \tag{5}$$

where $\varepsilon_{jt} = \epsilon_{jt} + \eta_{jt}$ and $\Phi_t(k_{jt}, i_{jt}) = \beta k_{jt} + \omega_{jt}$. Thus one can consistently estimate the labor coefficient α, the intermediate input coefficient γ, and the value $\Phi_{jt} = \Phi_t(k_{jt}, i_{jt})$ in (5) by semiparametric techniques (e.g., Robinson [1988]). This constitutes the first stage of the estimation. For the second stage, given the first stage estimates $(\hat{\alpha}, \hat{\gamma}, \hat{\Phi})$, and any possible value of the capital coefficient β, we can form the residual

$$\zeta_{jt}(\beta) = y - \hat{\alpha} l_{jt} - \hat{\gamma} m_{jt} - \beta k_{jt} = \omega_{jt} + \varepsilon_{jt} \tag{6}$$

for all observations (j, t). We can also form the residual

$$\omega_{jt}(\beta) = \hat{\Phi}_{jt} - \beta k_{jt} \tag{7}$$

for all observations (j, t) such that investment demand is positive, i.e., $i_{jt} > 0$. Then the nonparametric regression of ζ_{jt} on ω_{jt-1} yields

$$E[\omega_{jt} + \varepsilon_{jt} \,|\, \omega_{jt-1}, i_{jt-1} > 0] = E[\omega_{jt} \,|\, \omega_{jt-1}],$$

where the last equality follows from the fact that the mean zero ε_{jt} is independent of the firm's information set $\Im_{jt-1}$, and that ω_{jt} is independent of $\Im_{jt-1}$ conditional on ω_{jt-1} by the first order Markov assumption. Finally the parameter β is estimated using the moment condition that the innovation in period t's productivity, namely $\xi_{jt} = \omega_{jt} - E[\omega_{jt} \,|\, \omega_{jt-1}]$, is mean independent of k_{jt} since k_{jt} is an element of $\Im_{jt-1}$.

We end this description of the mechanics of OP with a brief cautionary note. The economic relationships in the model give rise to a number of alternative and seemingly more efficient ways of constructing moments. For example Ackerberg, Benkard, Berry, and Pakes [2007] advocate an "equivalent" method of constructing the moments. To quote the authors (p. 53, adjusting their notation to square with ours)

> An equivalent way to construct a moment condition ... is as follows ... Given β, construct $\hat{\omega}_{jt} = \hat{\Phi}_{jt} - \alpha k_{jt}$. Nonparametrically regress $\hat{\omega}_{jt}$ on $\hat{\omega}_{jt-1}$ to construct the estimated residual $\hat{\xi}_{jt}$. Construct a moment condition interacting $\hat{\xi}_{jt}$ with k_{jt}.

However this manner of constructing the moment suffers from a potential selection bias. To see this, recall (7) and notice that we can only construct the residual $\hat{\omega}_{jt}$ for those observations (j, t) such that $i_{j,t} > 0$. However

$$E[\omega_{jt} \,|\, \omega_{jt-1}, i_{jt} > 0, i_{jt-1} > 0] = E[\omega_{jt} \,|\, \omega_{jt-1}, i_{jt} > 0] \neq E[\omega_{jt} \,|\, \omega_{jt-1}].$$

Since a significant fraction of observations in typical plant level data sets have zero investment levels, this selection problem could be particularly problematic.

4 The LP Approach

Motivated by the fact that a large number of observations in typical plant level data report zero investment, LP sought an alternative to investment as the proxy variable used in the first stage of the OP procedure as a way to enhance the efficiency of the estimator. An important contribution of LP was to focus the analysis on the "inner" profit maximization problem (3) that the firm faces in the spot market, from which they show that intermediate inputs can serve as an alternative proxy variable. In particular, they prove that for a perfectly competitive firm, the input demand equation $m_{jt} = g_t(k_{jt}, \omega_{jt})$ is invertible in ω_{jt} under very general assumptions on the production technology.

One advantage of restricting the analysis to the static problem (3) rather than the full dynamic problem (4) is that it softens the bite of the scalar unobservability assumption: ω_{jt} is assumed to be the only firm specific unobserved state variable for just the spot market game rather than the whole dynamic game. Another advantage of working with (3) (emphasized by LP themselves) is that it is a simpler problem, and hence the challenge of proving invertibility of $g_t(k_{jt}, \omega_{jt})$ is made easier and perhaps less sensitive to the regularity conditions required by OP. We take their point a step further. As we will show, given the functional form of the production function itself, there is much more information contained in the firm's spot market problem than just the invertibility of $g_t(k_{jt}, \omega_{jt})$. Furthermore, we can exploit this information to completely relax the scalar unobservability assumption and thus generalize the LP approach to handle more complicated and unspecified patterns of unobserved heterogeneity among firms in the model.

5 Scalar Unobservability

The first key is to recognize that scalar unobservability remains a strong assumption even under the LP approach, and that it imposes a number of restrictions on the underlying economic model. We enumerate just a few of these restrictions below.

- It requires that all firms compete in the same input market. This, for example, assumes that firm's employ the same quality labor as one another (otherwise labor quality would be a firm specific state variable).
- If labor was a dynamic input, it requires that the costs of adjusting labor are identical for all firms regardless of their size (ignoring the problem that larger plants have possibly more unionized workers). More generally, all firms are assumed to face the same factor prices.
- All firms compete in the same output market, regardless of geography.
- The TFP process that describes a firm's evolution of productivity cannot have a firm

fixed effect (such as in the dynamic panel literature as discussed by Ackerberg et al. [2006]).

Most importantly for our purposes, the possible forms of product market competition that are consistent with the scalar unobservability assumption in (3) comprises a very short list. This is because the degree of product differentiation is being severely restricted by the assumption that the reduced form profit function has the form $\pi(k_{jt}, \omega_{jt}, \Delta_t)$. For example, in the usual Pakes and Ericson [1995] model of firm and industry dynamics, one would typically take the common state Δ_t to be the distribution over the firm specific states (k_{jt}, ω_{jt}) for all firms in the industry. Thus a firm's own profit depends upon its own state and the distribution over states of its competitors. But in the present context, since ω_{jt} is a scalar, there are no more firm specific state variables left to possibly to characterize the type of product that the firm produces. Thus all possible forms of product differentiation among firms has to be captured exclusively by the common state Δ_t, which essentially constrains product market competition between firms to fall into two possible cases:

1. *Homogeneous Products/Perfect Competition* : One possibility is that there is no product differentiation among firms. In light of the restriction that all firms compete in the same output market, and the fact that the asymptotics in the model require that the number of firms approach infinity (and within any sample, the data consist of usually at least a few hundred firms), this becomes the classic textbook case of a perfectly competitive market (for a more in depth view of limit theorems for homogeneous good industries approaching perfect competition with the number of firms, see e.g., Novshek and Chowdhury [2003]). In this case each firm acts as a price taker in the output market, and scalar unobservability would imply that the output price p_t is common across firms in a given time period.

2. *Symmetric differentiation/Monopolistic Competition* : To avoid the implication that the market approaches perfect competition as more firms enter, we need the product

space to not "fill up" with the number of firms. A necessary condition is that there exists persistent product differentiation among firms as more firms enter. Furthermore, due to the scalar unobservability assumption, this differentiation must be symmetric in nature, i.e., all firms are differentiated from each other in the same way so that the degree of competition can be captured by the common term Δ_t. This is satisfied, for example, through models of monopolistic competition with CES or logit type demand, in which case Δ_t might consist of the CES parameter that describes substitutability between firms.

Our approach is simply to use the power of the restrictions 1-2 on the nature of product market competition among firms implied by the scalar unobservability assumption rather than the scalar unobservability assumption itself. By assuming that product market competition takes either the form of 1 or 2, we can relax the firm's spot market profits to be of the form $\pi(l_{jt}, m_{jt}, k_{jt}, \Delta_{jt})$ rather than $\pi(l_{jt}, m_{jt}, k_{jt}, \Delta_t)$, with Δ_{jt} being an arbitrarily complicated unobserved state.

We do require one further weak assumption in the illustrations that follow below, which is that there exists at least one input that is competitive for all firms, i.e., all firms act as price takers in this input market (although firms need not face the same input prices). Here we draw upon the usefulness of the intermediate inputs used by LP, as energy or materials serve as natural candidates. Energy is particularly attractive as an input that can freely adjust each period with the firm acting as a price taker in the input market. Notationally, since we are using the single variable m_{jt} to generically represent all intermediate inputs, we let m_{jt} play the role of the competitive input.

We separately consider the cases of homogeneous products/perfect competition and symmetric differentiation/monopolistic competition. We emphasize the primal of the profit maximization problem in the former case, and the dual (cost minimization) in the latter case, to arrive at our empirical strategy. Nevertheless the estimating equations are highly related in both settings.

6 Homogeneous Products/Perfect Competition

The main idea is that under perfect competition in the product market, we can combine the first order condition for the competitive input with the production function itself and use both equations *jointly* to invert out productivity. Thus the first order condition for the intermediate input plays a role analogous to the proxy demand equation. However it contains much more information than the proxy inversion, which allows us to control for potentially complicated patterns of unobserved heterogeneity in input and output prices, adjustment costs, etc.

Recall the empirical specification (1) and the decomposition of the error term (2). These together imply a production function

$$Q_{jt} = A_{jt} U_{jt} L_{jt}^{\alpha} K_{jt}^{\beta} M_{jt}^{\gamma} \tag{8}$$

where $\log(A_{jt}) = \omega_{jt}$ is TFP and $\log(U_{jt}) = \epsilon_{jt}$ is the unanticipated productivity shock. Here we have merely backed out the underlying production function implied by the standard empirical specification. Recall $U_{jt} > 0$ is a productivity shock that the firm can only observe *after* it has made its period t input decisions. Thus when planning for these period t input decisions, the firm faces a *stochastic* production function in (8). The OP/LP model implies that the unanticipated productivity shock U_{jt} is independent of the productive inputs $(A_{jt}, L_{jt}, K_{jt}, M_{jt})$, i.e., it is an exogenous shock. Furthermore we assume that the firm has rational expectations and hence knows the distribution G_t from which the shock is drawn. This implies that WLOG, $E[U_{jt}] = 1$, since any systematic expectation for the unexpected shock other than one simply gets subsumed in TFP component A_{jt} from the point of view of the firm. Thus the firm can form the conditional expectation

$$Q_{jt}^{*} = E[Y_{jt} \mid A_{jt}, L_{jt}, K_{jt}] = A_{jt} L_{jt}^{\alpha} K_{jt}^{\beta}.$$

Finally observe that by definition the observable output Y_{jt} is related to the firm's expected output Q^*_{jt} by

$$y_{jt} = q^*_{jt} + \epsilon_{jt} + \eta_{jt}, \tag{9}$$

where recall lowercase letters denote logs.

We now consider the firm's intermediate input problem. Given potentially heterogeneous output prices P_{jt} and prices of the intermediate input W_{jt}, and assuming risk neutrality on the part of the firm, the first order condition for the intermediate input becomes.[2]

$$\gamma \frac{P_{jt} Q^*_{jt}}{M_{jt}} = W_{jt}. \tag{10}$$

Now taking logs of both sides, adding $\epsilon_{jt} + \eta_{jt}$ to both sides, rearranging terms and recalling the definition $\varepsilon_{jt} = \epsilon_{jt} + \eta_{jt}$, we get

$$\ln \left(\frac{W_{jt} M_{jt}}{P_{jt} Y_{jt}} \right) = \log(\gamma) - \varepsilon_{jt}. \tag{11}$$

Equation (11) provides the essential additional information to solve the endogeneity problem inherent in the production function (1). To see this, combine (1) and (11) to yield the system

$$\begin{aligned} \ln \left(\frac{W_{jt} M_{jt}}{P_{jt} Y_{jt}} \right) &= \log(\gamma) - \varepsilon_{jt} \\ y_{jt} &= \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + \omega_{jt} + \varepsilon_{jt}. \end{aligned} \tag{12}$$

Letting $s_{jt} = \ln \left(\frac{W_{jt} M_{jt}}{P_{jt} Y_{jt}} \right)$ and $x_{jt} = (l_{jt}, k_{jt}, m_{jt})$, we can express the above system more generally as

$$\begin{pmatrix} s_{jt} \\ y_{jt} \end{pmatrix} = \Upsilon(x_{jt}, \omega_{jt}, \varepsilon_{jt}). \tag{13}$$

[2] This FOC for the intermediate inputs holds regardless of whether labor is static or dynamic, and fixed or variable. In the event that labor is dynamic, the FOC holds regardless of the presence of an arbitrary firm and time specific cost of adjusting labor, i.e., $c_{jt}(l_{jt}, l_{jt-1})$. This flexibility would not be permitted under scalar unobservability, which would require $c_{jt}(l_{jt}, l_{jt-1}) = c(l_{jt}, l_{jt-1}, \Delta_t)$.

For production functions more general than Cobb-Douglas, the system that results from combining the first order condition for labor with the production function will yield the form (13). Indeed, for a variety of production functions that we have tried, the system in fact takes the simple triangular form in (17).

Now notice that for any realization of the data $D_{jt} = (s_{jt}, y_{jt}, l_{jt}, k_{jt})$ and value of the parameter vector (β, α), we can uniquely solve the two equation system (13) for the two econometric unobservables $(\omega_{jt}, \varepsilon_{jt})$. This is especially clear in the case of the system in (17), which forms a simple triangular system in $(\omega_{jt}, \varepsilon_{jt})$. In a later section, we show how the inversion is carried out with a CES production function, which also forms a triangular system.

To estimate the parameter vector, we simply use the variety of possible moment conditions offered by the underlying economic model, which are exactly those used by OP and LP. In particular, we can implement similar moments to those suggested by Ackerberg et al. [2006]. For example, the moment $E[\varepsilon_{jt}] = 0$ identifies γ. Furthermore, the fact that innovation ξ_{jt} is independent of the firm's information set at time $t-1$ implies $E[\xi_{jt} k_{jt}] = 0$ (recall that innovation is $\xi_{jt} = \omega_{jt} - E[\omega_{jt} \mid \omega_{jt-1}]$), which identifies β. If one wishes to assume that labor is a variable input, then $E[\xi_{jt} l_{jt-1}] = 0$ identifies α. On the other hand, if one wishes to let l_{jt} be a fixed state variable that accumulates with a "time to build" assumption in the same fashion as capital (for example, hiring/firing decision may take time to enact and thus today's labor is effectively decided by a previous period's hiring/firing decisions), then $E[\xi_{jt} l_{jt}] = 0$ identifies α. In this case, the researcher can still use $E[\xi_{jt} l_{jt-1}] = 0$ as an overidentifying restriction.

Several points are to be noted concerning our approach. First, the role of introducing the first order condition for the intermediate input, as the system (13) makes clear, is to separate out the error terms ω_{jt} and ε_{jt}, which cannot be separated out from the production function alone. The intuition for why this separation is possible in a general production function setting is straightforward. The firm observes ω_{jt} and thus it gets absorbed in the first order condition, which is nothing other than usual endogeneity problem. However the expectational error/unanticipated shock ε_{jt} does not get absorbed by the firm but rather gets introduced when we move from the first order condition from the point of view of the firm to the first order condition from the point of view of the econometrician, i.e., the transition from (10) to (11). Thus the two error terms enter asymmetrically into the second equation, which is the key to the ability to invert the system (17). In a recent paper, Ackerberg et al. [2006] modify the proxy variable approach of OP and LP so that the first stage regression has a similar function to our first order condition, namely to break the symmetry between ω_{jt} and ε_{jt} in the production function.

Another key point is that we allow for input and output price heterogeneity. However it is not required that we be able to observe these prices since they enter the second equation through the term s_{jt}, and hence we only need to observe the firm's expenditure on the intermediate input $W_{jt}M_{jt}$ and its revenue $P_{jt}Y_{jt}$. Both of these expenditure variables are readily available in typical plant level data sets. In fact, the measure of output Y_{jt} in these data sets is typically derived by deflating revenue $P_{jt}Y_{jt}$, where revenue is what the firm actually reports to the census authority (we return to the potential problem caused by revenue production functions in a later section). The firm's expenditure on inputs is also typically what is reported, and hence measuring M_{jt} similarly requires deflating $W_{jt}M_{jt}$. Thus $W_{jt}M_{jt}$ and $P_{jt}Y_{jt}$ are actual primitives in the data that do not suffer from measurement error generated by applying deflators (such as is required when trying to measure real intermediate input demand or real investment when applying the proxy, which require deflating expenditures on intermediate inputs and investment).

7 How We Solved the Colinearity Problem?

A major question that has recently been raised about the proxy variable approach, most notably by Bond and Söderbom [2005] and Ackerberg et al. [2006] is how the model is capable of identifying the coefficient on an input that is both variable and static input. In short, the endogeneity problem problem in the production function generates a colinearity problem among the inputs that stands as an obstacle to any method for identifying production function parameters. The most natural "way out" of the colinearity problem is to assume the existence of input price variation between firms, but such variation is ruled out by the assumptions of OP/LP model. As we explained in the previous section, the existence of such unobserved heterogeneity in input prices and other factors was one of our motivations for pursuing the identification strategy based on the system (17), and is thus one reason we are able to solve the colinearity problem. However as we explain in this section, input price variation is not required by our methodology, and the colinearity problem in our setting is being solved more fundamentally by our use of information contained in the firm's profit maximization problem.

The colinearity problem results from the fact that any input in the production function that is both static and variable (we shall simply refer to these as the VS inputs) can be expressed as a function of the fixed inputs and the firm's TFP. If this functional relationship does not vary across firms in the population (for example, due to input price variation), then this causes a basic identification problem to arise that breaks down the first stage of the OP/LP procedures. We illustrate this problem with a simple two input production function, one input v being the VS input, the other input k being the fixed input, and d being the proxy variable used by OP/LP to control for the endogeneity problem. Thus the production function for firm j in period t is

$$y_{jt} = \alpha v_{jt} + \beta k_{jt} + \omega_{jt} + \varepsilon_{jt}.$$

Recalling the scalar unobservable assumption, we have that demand for the VS input can be expressed as $v_{jt} = f_t(k_{jt}, \omega_{jt})$. However the proxy demand also has the form $d_{jt} =$

$\delta_t(k_{jt}, \omega_{jt})$, which we invert to obtain $\omega_{jt} = h_t(k_{jt}, d_{jt})$. From this it follows that

$$v_{jt} = f_t(k_{jt}, h_t(k_{jt}, d_{jt})) = F_t(k_{jt}, d_{jt}).$$

Going back to the first stage regression in the proxy variable approach, if we wish to be fully nonparametric about the control function term Φ_t in (5), then it follows that we cannot simultaneously identify the labor coefficient α. To see this, consider $\alpha' \neq \alpha$. Then

$$\begin{aligned} y_{it} &= \alpha v_{jt} + \Phi_t(k_{jt}, d_{jt}) + \varepsilon_{jt} \\ &= (\alpha' + (\alpha - \alpha'))v_{jt} + \Phi_t(k_{jt}, d_{jt}) + \varepsilon_{jt} \\ &= \alpha' v_{jt} + (\alpha - \alpha')F_t(k_{jt}, d_{jt}) + \Phi_t(k_{jt}, d_{jt}) + \varepsilon_{jt} \\ &= \alpha' v_{jt} + \Phi_t'(k_{jt}, d_{jt}) + \varepsilon_{jt}. \end{aligned}$$

Thus we cannot be both fully nonparametric about Φ_t and still identify α. As Bond and Söderbom [2005] and Ackerberg et al. [2006] show, the same situation persists even if we use the parametric structure on Φ_t implied by the Cobb-Douglas specification and perfect competition.

How does the identification strategy that we presented in Section (6) solve the colinearity problem? Recall that the essence of our approach was to use the FOC with respect to the VS input as a second estimating equation. The endogenous part of the error term, namely ω_{jt}, gets "internalized" in the profit maximizing level of the variable input's revenue share s_{jt}. However s_{jt} is the dependent variable in our second equation, and hence there is no endogeneity problem (and as a result, no colinearlity problem) in the second equation taken by itself. Thus we are able to identify the coefficient on the VS input since our second equation provides a source of information on this coefficient that is free of endogeneity and colinearity concerns. Of course, if our only problem was to estimate the VS coefficient, then we could use the FOC by itself to estimate α. However to identify the full set of production function parameters, we use the full system (13). This same logic extends readily to other functional forms as a result of the triangular structure that (13) takes.

8 Empirical Example

We now apply our method to the same Chilean manufacturing data used in Levinsohn and Petrin [2003]. In particular, the data is identical to that employed by Greenstreet [2007], which employs the latest "cleaning" technologies (the appendix to Greenstreet [2007] contains an extensive description). We focus on the largest industry, namely 311 (food products). We employ a three input production function in energy, labor, and capital, and use real value added (doubly-deflated) constructed for these three inputs as our output measure. Energy serves as the intermediate input entering our FOC.

To highlight the possibly complicated nature of the labor input problem, we model labor as a fixed state variable that accumulates in a time to build fashion. In order to compare our estimator (hereafter abbreviated as GNR) against the proxy variable approach, we must use an analogue of the OP/LP estimator proposed by Ackerberg et al. [2006] (hereafter abbreviated as ACF) that allows labor to be a fixed state variable. We use the energy input as the proxy variable for ACF and use it as the input to the first order condition for GNR. For both estimators, we use the moment $E[\xi_{jt}l_{jt}] = 0$ to identify the labor coefficient. Finally we employ bootstrap standard errors for both estimators. The results are as follows.

Table 1: Industry 311: Cobb-Douglas

Method	Labor	SE	Capital	SE	Energy	SE
OLS	.608	.009	.210	.006	.349	.009
ACF	.483	.039	.160	.021	.510	.054
GNR	.659	.025	.315	.016	.087	.001

The most apparent feature that emerges from the table is the extent to which the estimates obtained under GNR reduce the energy coefficient as compared to both OLS and ACF. The comparison between GNR and OLS can easily be understood in light of the simultaneity problem. Since energy in the model is the flexible input, this will translate into an upwardly biased coefficient for energy in OLS. Our FOC is correcting this upward bias. Furthermore, as discussed in depth by LP, the coefficients on the quasi-fixed inputs, which in this case are labor and capital, will generally be biased downward as a result of the coefficient on the flexible input being biased upward. We see this playing out in our results, with GNR producing a larger capital and labor coefficient as compared to OLS. Finally, the GNR estimator yields a returns to scale not significantly different than 1. It is not well understood at this time what is driving the divergence between ACF and GNR - the main observation to make is that ACF is pulling the coefficients in the opposite direction relative to OLS as compared to GNR (i.e., labor and capital coefficients go down, but energy goes up).

9 Checking Robustness : The CES production function

We now show how to invert a CES production function and carry out our empirical strategy. In particular, we are naturally concerned that our above findings are possibly tied to the particular functional form of a Cobb-Douglas. Other functional forms we have tried include a Translog and Diewert production function, which we do not exposit here. The CES is of course a natural alternative to try since it is well known that the Cobb-Douglas and the CES functional form exhaust the space of production functions that are homogeneous and additively separable. In what follows below, we change notation slightly since we have settled on energy as our intermediate input. Thus let N denote the quantity of energy, and δ denote the price of energy.

Take the following CES production function:

$$Q_{jt} = A_{jt} \left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho} \right)^{\frac{r}{\rho}}$$

Let small letters represent logarithms and let $\ln(A_{jt}) = \omega_{jt}$. Further, let $Y_{jt} = Q_{jt}e^{\varepsilon_{jt}}$ denote measured output where ε_{jt} is the mean zero exogenous error.

The expected profit function will be

$$P_{jt}E(Q_{jt}) - c_{jt}(N_{jt}, K_{jt}, L_{jt}) = P_{jt}A_{jt}\left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho}\right)^{\frac{r}{\rho}} - \delta_{jt}N_{jt} - c_{jt}(K_{jt}, L_{jt}).$$

If we assume that energy (N_{jt}) is bought by the firm in a competitive market (or at least that the firm behaves as a price taker) the FOC for energy (N_{jt}) is:

$$\begin{aligned}
\frac{r}{\rho}P_{jt}A_{jt}\left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho}\right)^{\frac{r}{\rho}-1}\gamma\rho N_{jt}^{\rho-1} - \delta_{jt} &= 0 \iff \\
\gamma r P_{jt}A_{jt}\left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho}\right)^{\frac{r}{\rho}-1}N_{jt}^{\rho} &= \delta_{jt}N_{jt} \iff \\
\gamma r P_{jt}Q_{jt}\left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho}\right)^{-1}N_{jt}^{\rho} &= \delta_{jt}N_{jt} \iff
\end{aligned}$$

$$\frac{P_{jt}Q_{jt}}{\delta_{jt}N_{jt}} = \frac{1}{\gamma r}N_{jt}^{-\rho}\left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho}\right)$$

and taking logs and writing in terms of observed variables

$$\ln\left(\frac{P_{jt}Y_{jt}}{\delta_{jt}N_{jt}}\right) = -\ln(\gamma r) - \rho n_{jt} + \ln\left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho}\right) + \epsilon_{jt} \tag{14}$$

which, with the production function

$$y_{jt} = \frac{r}{\rho}\ln\left(\alpha L_{jt}^{\rho} + \beta K_{jt}^{\rho} + \gamma N_{jt}^{\rho}\right) + \omega_{jt} + \epsilon_{jt}$$

gives us a system of equations on the two unobservables $(\omega_{jt}, \varepsilon_{jt})$.

We now estimate the three input production function for industry 311 using the GNR approach. First observe however that the CES does not impose that the elasticity of labor and capital are constant (in contrast to Cobb-Douglas), but allows the elasticity to vary across firms depending on the chosen input levels. Thus we report the average elasticity of labor and capital among firms in the data to see if they roughly square with what we found for the Cobb-Douglas case. We find the following.

Table 2: Industry 311: CES

	Estimate	Std. Err.
Mean Capital Elasticity	0.373	0.008
Mean Labor Elasticity	0.493	0.024
Mean Energy Elasticity	0.118	0.002
Returns to Scale	0.984	0.013
Inverse Elas. of Substitution	0.960	0.008

Our results produce elasticities in the same "ballpark" as Cobb-Douglas. The average labor elasticity become smaller and the average capital elasticity becomes bigger, with the average energy elasticity remaining roughly the same. Moreover returns to scale remains roughly 1, and the CES parameter is close (but significantly different than) 1, which means we find a small departure from Cobb-Douglas.

The fact that the CES allows for heterogeneity in input elasticities leads one to ask whether the fact that these elasticities are held fixed under Cobb-Douglas has any testable implications. We pursue such a testable implication in the next section.

10 Nonparametric Implications of Cobb-Douglas.

We now consider whether our base model - a Cobb-Douglas production function with fixed coefficients and perfectly competitive firms in the product market who face at least one competitive input - has any testable implications. We derive one such implication by way of the first order condition for the intermediate inputs. In particular, we ask what would be implied of the data if the coefficient on energy, γ, was a random coefficient. That is, each firm j would have its own γ_j, and the distribution of coefficients in the population is some unknown G.

Consider a version of equation (11), where γ is replaced with γ_j which varies across firms j. In particular assume each γ_j is drawn from a distribution G. Assume each firm is in operation for at least two time periods which we generically call $t = 1, 2$. Consider now the system (11) for periods $t = 1$ and $t = 2$, which yields

$$\ln(\frac{\delta_{j1} N_{j1}}{P_{j1} Y_{j1}}) = \ln(\gamma_j) - \varepsilon_{j1}$$
$$\ln(\frac{\delta_{j2} N_{j2}}{P_{j2} Y_{j,2}}) = \ln(\gamma_j) - \varepsilon_{j2}. \tag{15}$$

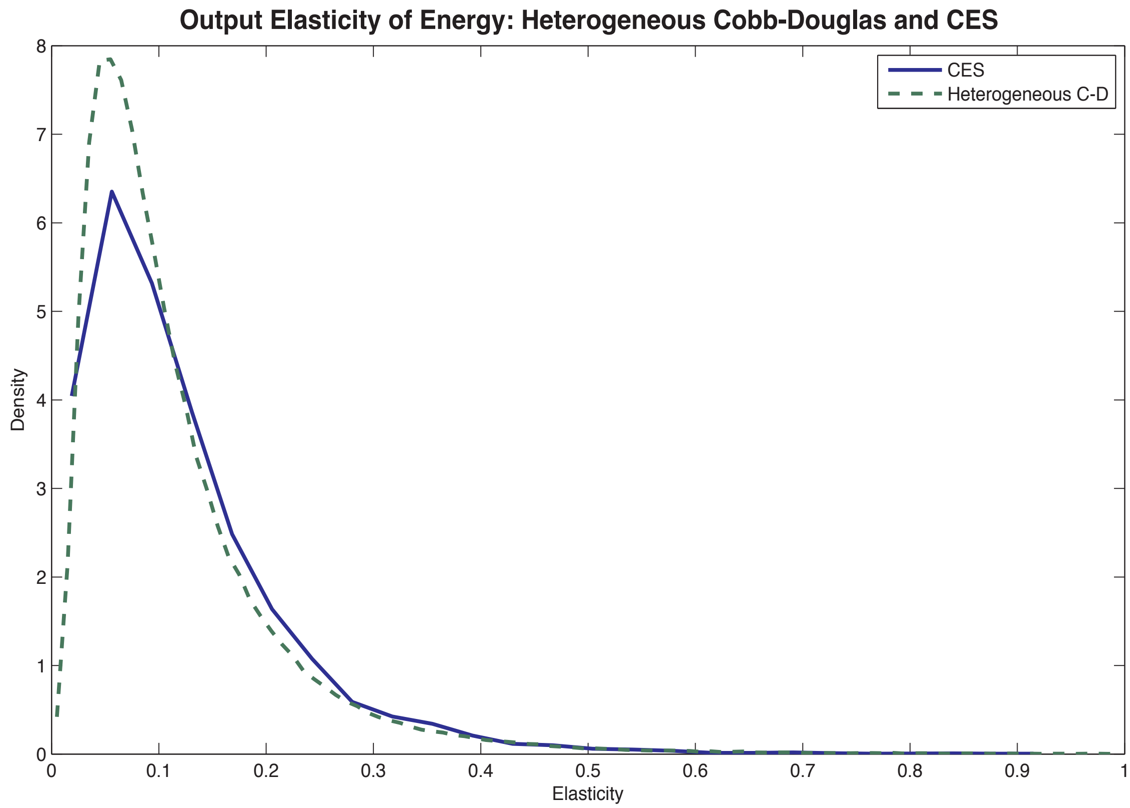
Let $s_{j1} = \ln(\frac{\delta_{j1} N_{j1}}{P_{j1} Y_{j1}})$ and $s_{j2} = \ln(\frac{\delta_{j2} N_{j2}}{P_{j2} Y_{j2}})$. From the data, we can directly recover the joint distribution of the random vector (s_1, s_2). As we show below it follows that, from a theorem of Kotlarski [1967], we can separately identify *nonparametrically* the distribution of the random variable $\gamma \sim G$ and the distributions of the random variables $\varepsilon_i \sim F_i$ for $i = 1, 2$.

Theorem 1. *If the joint distribution of (s_1, s_2) admits a non-vanishing characteristic function and the random variables $(\gamma, \varepsilon_1, \varepsilon_2)$ are mutually independent, then the joint distribution of (s_1, s_2) identifies the the joint distribution of $(\gamma, \varepsilon_1, \varepsilon_2)$.*

Proof. Recall the assumption that $E[\varepsilon_i] = 0$ for $i = 1, 2$. With this restriction in place, refer to the system (15). The theorem of Kotlarski [1967] implies that that the distribution of $g = \ln(\gamma)$ and the distributions of μ_i for $i = 1, 2$ are nonparametrically identified from the joint distribution of (s_1, s_2). Thus we can identify the distribution G of the one-to-one transformation $\exp(g) = \gamma$ and the distribution F_i of ε_i for $i = 1, 2$. $\square$

The basic idea behind the identification theorem is that in a single cross section (for say $t = 1$), we cannot separately identify whether the heterogeneity in the intermediate input's share of revenue across firms is due to heterogeneity in γ, or heterogeneity in the shocks ε_1. However, if we bring a second cross section to bear on the problem through $t = 2$, we can use the *persistence* in the input's share of revenue within a firm across time to separate out the effects. Thus the higher the correlation in the input's share of revenue between the two time periods, the more of the observed cross sectional variance in the input's share of revenue we attribute to differences in γ.

To implement the estimation of the the distribution of $(\gamma, \{\varepsilon_t\}_{t=1}^{T})$ in the system (15), we can apply semiparametric factor methods such as those described by e.g., Carneiro et al. [2003]. We now present estimates of the distribution of γ for industry 311 from the Chilean data (which we used in Section 7) assuming that each period t's exogenous shock ε_t comes from a common distribution F (an assumption that our identification theorem shows that we can relax). While we can recover the distributions nonparametrically, we instead estimate the distributions F and G using flexible parametric forms. In particular, we let the distribution of $g = \gamma$ follow a truncated (between $(0, \infty)$) mixture of normals with 5 components and F follow a mixture of normals with 3 components and obtain estimates by maximum likelihood. In both cases we treat missing observations for a firm as missing at random.



In the above graph, we present our estimated distribution of γ against the distribution of the energy elasticity from the CES estimates. Interestingly, the distribution we recover from the de-convolution method matches very closely to the sample distribution in the energy elasticity over firms that was estimated for CES.[3] Thus it appears that in so far as heterogeneity in the energy coefficient is playing a role, it is being captured well by the more flexible CES functional form. Essentially, the CES is explaining the perisistence of the input's revenue share over time because of persistence in the firm's input choices over time (as is evident from (14)). Understanding further the role that CES functional form in explaining this pattern in the data requires further research.

11 Product Differentiation with Quantity Data

11.1 Symmetric Differentiation/Monopolistic Competition

In this section we relax the assumption that firms face a perfectly competitive environment in the product market. In order to do so, we work with the dual of the problem and assume that firms are cost minimizers instead. We keep the assumption that at least one input (the intermediate input M_{jt}) is chosen every period in a competitive market.

Firm $j's$ problem with respect to the variable static intermediate input M_{jt} will be

$$\min_{M_{jt}} w_{jt}M_{jt} + \lambda_{jt}\left(Q^* - A_{jt}U_{jt}L_{jt}^{\alpha}K_{jt}^{\beta}M_{jt}^{\gamma}\right)$$

where, from standard envelope arguments, it follows that the multiplier λ_{jt} is the marginal cost of firm j. In this case, the FOC for the intermediate input implies that

$$\ln\frac{w_{jt}M_{jt}}{P_{jt}Y_{jt}} = -\theta_t + \ln\gamma - \varepsilon_{jt} \tag{16}$$

where, because we are assuming symmetric differentiation/monopolistic competition, the log markup $\theta_t = \ln\frac{P_{jt}}{\lambda_{jt}}$ is constant across firms.[4]

To understand how equation (16) helps, first assume that we can observe Y_{jt} directly. In

[3]In an appendix we show how to estimate the other parameters of the production function when there are firm specific γ's.

[4]The assumption that the markup is constant can be relaxed using the arguments in section 9.

this case, we have

$$y_{jt} = \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + \omega_{jt} + \varepsilon_{jt}$$

which coupled with equation (16) generates a triangular system of equations as before. Notice that now, however, $E(\varepsilon_{jt}) = 0$ is not enough to identify γ and we would need to use a moment like $E(\xi_{jt} m_{jt-1}) = 0$ to recover both γ and θ_t.

11.2 Firm Specific Product Differentiation

In this section we show that by adding the information contained in the proxy demand equation we can estimate firm and time specific markups. As we show below, when quantities are observed we can recover these markups without additional assumptions about the demand facing the firm. Furthermore, as we show in the next section, when only revenues are observed, firm and time specific markups can still be recovered by assuming the standard CES demand system.

Suppose that there exists one static input which we will call M_{jt}, and one proxy variable ι_{jt}. Our method proposes using the following system of equations to estimate the production function.

$$\ln\left(\frac{W_{jt}M_{jt}}{P_{jt}Y_{jt}}\right) = -\theta_{jt} + \log(\gamma) - \varepsilon_{jt} \tag{17}$$

$$y_{jt} = \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + \omega_{jt} + \varepsilon_{jt}$$

If we add to this system a modified version of the proxy variable equation, then this additional source of information allows us to identify the firm and time varying markups. Now, the demand for the proxy variable will be a function of the markup, θ_{jt}.

$$\iota_{jt} = f_t(k_{jt}, l_{jt}, \omega_{jt}, \theta_{jt})$$

Notice that this setup is not possible under OP (or LP or ACF) due to the scalar unobservability assumption. Since the markup varies by firm, it cannot be controlled for by differences across time in $f_t\left(\right)$. It is possible here because the first order condition provides an additional source of information (*i.e.* it adds $J \times T$ equations to identify the $J \times T$ additional parameters).

This gives us a new system of equations.

$$
\begin{aligned}
\ln\left(\frac{W_{jt}M_{jt}}{P_{jt}Y_{jt}}\right) &= -\theta_{jt} + \log\left(\gamma\right) - \varepsilon_{jt} \\
y_{jt} &= \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + \omega_{jt} + \varepsilon_{jt} \\
\iota_{jt} &= f_t\left(k_{jt}, l_{jt}, \omega_{jt}, \theta_{jt}\right)
\end{aligned}
$$

that can be solved (for ω_{jt}, θ_{jt}, and ε_{jt}) for any value of the parameter vector (α, β, γ). Given the solution we can form moments on ξ_{jt} and ε_{jt} and set up the GMM problem as before.

[5]

12 Revenue Production Functions

An issue that has received some attention recently is the fact that often we do not observe output directly, and instead need to deflate revenues to obtain quantities. As first pointed out by Klette and Griliches [1996] and addressed by several other papers including Syverson (2007) and De Loecker [2007], when there are price differences between firms, then revenues will not be properly deflated, and the difference between a firm's price and the average industry price will appear in the error term of the production function. Because these price differences are likely correlated with the observed inputs, this generates an endogeneity problem. In order to control for these price differences, these papers introduce a demand system and jointly estimate it with the production function. We can do the same, however

[5] Alternatively, we can treat this as a constrained minimization problem (see Judd and Su [2008]).

we are not restricted to assuming constant markups and we can recover firm and time specific markups.

Suppose we observe deflated revenues, $R_{jt} = \frac{P_{jt}Q_{jt}}{P_t}$, instead of quantities. Then the production function takes the following form.

$$r_{jt} = \alpha l_{jt} + \beta k_{jt} + \omega_{jt} + \ln\left(\frac{P_{jt}}{P_t}\right) + \varepsilon_{jt} \tag{18}$$

The problem is that P_{jt} is unobserved and likely correlated with inputs. By bringing in a demand system we can control for these unobserved firm-level prices.

Let demand have the following form:

$$Q_{jt} = Q_t\left(\frac{P_{jt}}{P_t}\right)^{\frac{1}{\eta_{jt}}}$$

Then substituting for Q_{jt} using the production function and taking logs we can rewrite this as

$$\ln\left(\frac{P_{jt}}{P_t}\right) = \eta_{jt}\alpha l_{jt} + \eta_{jt}\beta k_{jt} + \eta_{jt}\omega_{jt} - \eta_{jt}q_t. \tag{19}$$

Replace equation (19) into (18) to obtain a new production function

$$r_{jt} = \alpha\left(1+\eta_{jt}\right)l_{jt} + \beta\left(1+\eta_{jt}\right)k_{jt} + \left(1+\eta_{jt}\right)\omega_{jt} - \eta_{jt}q_t + \varepsilon_{jt}. \tag{20}$$

Note that the markup, θ_{jt}, and the elasticity of demand, η_{jt}, are related through the equation, $\theta_{jt} = \frac{1}{1+\eta_{jt}}$. This gives us a revised version of our previous system of three equations and three unknowns.

$$\begin{aligned}
\ln\left(\frac{W_{jt}M_{jt}}{P_{jt}Y_{jt}}\right) &= -\theta_{jt} + \log\left(\gamma\right) - \varepsilon_{jt} \\
r_{jt} &= \alpha\left(\frac{1}{\theta_{jt}}\right)l_{jt} + \beta\left(\frac{1}{\theta_{jt}}\right)k_{jt} + \left(\frac{1}{\theta_{jt}}\right)\omega_{jt} - \left(\frac{1-\theta_{jt}}{\theta_{jt}}\right)q_t + \varepsilon_{jt} \\
\iota_{jt} &= f_t\left(k_{jt}, l_{jt}, \omega_{jt}, \theta_{jt}\right)
\end{aligned}$$

This system can be solved (for ω_{jt}, θ_{jt}, and ε_{jt}) and the parameters estimated in the same way as in Section (11.2). Also note that, as opposed to for example De Loecker [2007], we do not need variation in q_t to identify the markups, due to our additional equation, the first order condition. The fact that we can identify firm and time specific markups is a new result.

13 Conclusion

References

Daniel Ackerberg, Lanier Benkard, Steven Berry, and Ariel Pakes. Econometric tools for analyzing market outcomes. In James J. Heckman and Edward Leamer, editors, *Handbook of Econometrics*, volume 6. Elsevier Science, 2007. Forthcoming.

Daniel A. Ackerberg, Kevin Caves, and Garth Frazer. Structural identification of production functions. Unpublished Manuscript, UCLA Economics Department, 2006.

Eric J. Bartelsman and Mark Doms. Understanding productivity: lessons from longitudinal microdata. Finance and Economics Discussion Series 2000-19, Board of Governors of the Federal Reserve System (U.S.), 2000.

Stephen Bond and Måns Söderbom. Adjustment costs and the identification of cobb douglas production functions. Unpublished Manuscript, The Institute for Fiscal Studies, Working Paper Series No. 05/4, 2005.

Pedro Carneiro, Karsten Hansen, and James J. Heckman. Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review*, 44(2):361–422, May 2003. 2001 Lawrence R. Klein Lecture.

Jan De Loecker. Product differentiation, multi-product firms and estimating the impact of trade liberalization on productivity. NYU Stern, Unpublished Manuscript, 2007.

David Greenstreet. Exploiting sequential learning to estimate establishment-level productivity dynamics and decision rules. Economics Series Working Papers 345, University of Oxford, Department of Economics, 2007.

Zvi Griliches and Jacques Mairesse. Production functions: The search for identification. In *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, pages 169–203. Cambridge University Press, New York, 1998.

Kenneth L. Judd and Che-Lin Su. Constrained optimization approaches to estimation of structural models. Technical report, April 2008.

Tor Jacob Klette and Zvi Griliches. The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11(4): 343–361, July 1996.

Ignacy I. Kotlarski. On characterizing the gamma and normal distribution. *Pacific Journal of Mathematics*, 20:69–76, 1967.

James Levinsohn and Amil Petrin. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(243):317–342, April 2003.

Jacob Marschak and W.H. Andrews. Random simultaneous equations and the theory of production. *Econometrica*, 12:143–205, 1944.

William Novshek and Prabal Roy Chowdhury. Bertrand equilibria with entry: limit results. *Inernational Journal of Industrial Organization*, 21(6):795–808, 2003.

G. Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297, November 1996.

Ariel Pakes. The estimation of dynamic structural models: Problems and prospects, part ii. mixed continuous-discrete control models and market interactions. In J.J. Laffont and C. Sims, editors, *Advances in Econometrics: Proceedings of the 6th World Congress of the*

Econometric Society, chapter 5, pages 171–259. Cambridge University Press, New York, 1994.

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

Peter M. Robinson. Root-n-consistent semiparametric regression. *Econometrica*, 56(4):931–954, July 1988.