Levinsohn and Petrin's (2003) Methodology Works under Monopolistic Competition

Sergio DeSouza
Universidade Federal do Ceara

Abstract

Markups, returns to scale and productivity can be uncovered from regressing output on inputs. However, econometric identification of theses parameters may be problematic due the simultaneity problem. A common solution is the IV method. However, usual instruments are only weakly correlated to the explanatory variables. Levinsohn and Petrin (2003) propose using a commonly observable variable (intermediate input) to control for unobserved productivity. Their methodology is based on the following key result: under the assumption of perfect competition, the intermediate input's demand function is a monotonic function of productivity. However, firms in most industries enjoy some degree of market power such that perfect competition may not be a desirable assumption for most empirical studies. This paper contributes to the literature by showing the monotonicity condition holds under monopolistic competition.

I would like to thank FUNCAP for financial support.
Citation: DeSouza, Sergio, (2006) "Levinsohn and Petrin's (2003) Methodology Works under Monopolistic Competition." *Economics Bulletin,* Vol. 12, No. 6 pp. 1-11
Submitted: February 13, 2006. **Accepted:** August 9, 2006.
URL: http://economicsbulletin.vanderbilt.edu/2006/volume12/EB-06L10006A.pdf

1. Introduction

Identification of production functions has played an important role in applied economics. Indeed, meaningful parameters (markups, returns to scale and productivity) can be uncovered from a regression of output on inputs. The seminal contribution of Marschak and Andrews (1944) however points out that econometric identification of these parameters may be problematic due to the potential correlation between inputs and productivity. It is a straightforward exercise to demonstrate that this correlation yields biased OLS estimates.

A common solution to this problem is to treat productivity as a time-invariant term and use the within estimator. Unfortunately, the assumption of productivity as a fixed effect may be quite restrictive, especially for panels with a relatively large time dimension. Another approach is to difference the variables and use lagged inputs as instruments (Blundell and Bond, 2000). However, differencing removes much of the variation in the explanatory variables and instruments are only weakly correlated with the differenced explanatory variables (Wooldridge, 2005).

In the absence of good instruments applied economists started searching for alternative methods to deal with the simultaneity problem. Leading this recent literature Olley and Pakes (1996) -OP from now on- propose an innovative technique that avoids the difficult task of searching for instruments. They use an observed variable (investment) to proxy unobserved productivity. Another approach was developed by Levinsohn and Petrin (2000)-LP hereafter. They argue that the necessity to drop zero-investments observations, as required by the OP methodology, raises an estimation efficiency issue. Indeed, in commonly found data sets, especially those covering developing countries, the application of the investment proxy approach requires dropping many zero-investment data points. For instance, in the Chilean firm-level data set, used by LP, over one-half of the sample reports zero investment. This zero-investment problem may be evidence of a more fundamental issue. LP argues that if the zeros are the result of adjustment costs that lead to kink points in the demand function, plants may not entirely respond to productivity shocks such that the endogeneity problem can remain.

Following a similar approach but using a different variable to control for unobserved productivity LP devise an alternative framework. Their methodology can be briefly described as follows. First, they derive the intermediate input's demand function, which under certain assumptions is a monotonic function of productivity. Then, with the monotonicity condition at hand, they are able to invert the intermediate input's demand function to uncover the unobservable productivity term as a non-parametric function of the intermediate input and capital. In this way, the only unobservable error term left in the estimation is not expected to be correlated with the regressors. Note also that the LP technique is analytically much simpler than OP. Unlike investment, the intermediate input is a flexible variable. Then a simple static setup suffices to derive the monotonicity condition. In turn, OP have to incorporate dynamics to account for capital adjustments costs to prove that productivity is a monotonic function of investment. This is certainly not a trivial exercise.

In many data sets firms almost always report positive use of intermediate inputs. In such cases, the LP methodology becomes the natural choice. Although this methodology avoids some of the problems caused by the investment proxy it introduces a new one. As described in further detail in the following section, they assume perfect competition in order to derive the monotonicity condition. This assumption contrasts with the fact that firms in most industries enjoy some degree of market power. This paper contributes to the literature by relaxing the assumption of perfect competition and showing that the LP procedure can also be applied to study imperfectly competitive industries[1]. This paper is organized as follows. The next section briefly discusses the LP procedure and, more importantly, demonstrates the following result: aside from some regularity assumptions, the monotonicity condition holds under a particular form of imperfect competition. Section 3 shows an example that illustrates the usefulness of this result. Finally, the last section provides some concluding remarks and suggestions of future work.

2. Monotonicity Condition under Monopolistic Competition

In this section I demonstrate that the monotonicity condition is valid under the assumption of monopolistic competition. But first let me briefly explain how the LP methodology works.

Assume the inputs available to firm i are: freely variable inputs (labor, L_{it}, and intermediate input, M_{it}) and a quasi-fixed input (capital, K_{it}). These inputs are transformed into output (Q_{it}) according to the following production function.

$$Q_{it} = F(L_{it}, K_{it}, M_{it}, W_{it}, N_{it}, \Theta) \tag{1}$$

where Θ is a set of parameters and the error term has two additively separable components: W_{it} and N_{it}. The former term is the productivity shock observed by firms before they choose optimal labor and intermediate input levels, while the latter is an i.i.d random shock[2].

---

[1] Ackerberg et al. (2005) point out another problem with the LP approach. They argue that in the same way intermediate input is a function of productivity so is labor. Then, in a typical production function regression where the variable inputs appear on the RHS a colinearity problem arises, casting doubt on the parameters estimation. Further, they develop a method does not suffer from this colinearity problem, but they also use the monotonicity condition in (2) to control for unobservable productivity. Thus, this paper's result also broadens the set of assumptions on competition under which their methodology works.

[2] The first term is a state variable affecting firm's decisions while the second term has no impact on firm's controls. Olley and Pakes (1996) interpret the N_{it} as a shock to productivity that is unobserved by firms during the period in which the flexible inputs levels are optimized.

The basic idea behind LP is to proxy unobserved productivity with an observable variable. One candidate for such proxy is M_{it}. Note that the intermediate input's demand function is given as[3]

$$M_{it} = M_{it}(W_{it}, K_{it}) \tag{2}$$

If this function is strictly monotonic in W_{it} one can invert it, conditional on capital, i.e. $W_{it} = W_{it}(M_{it}, K_{it})$. This is the key result that supports the LP framework. Indeed, if monotonicity holds one can substitute W_{it} for $W_{it}(M_{it}, K_{it})$ in (1). Thus, the production function regression contains only observables and an error term, N_{it}, which is not correlated with the inputs. Then, LP devise a two-step algorithm to estimate the parameters. For more details see their original work. In this work, I am rather interested in the assumptions that support the monotonicity result and the limitations they may impose on empirical studies. For notational simplicity I will drop the time subscripts from now on.

LP show that the intermediate input demand function is strictly increasing in W_i, conditional on K_i, if the following assumptions hold:

> **Assumption 1**. Capital is quasi-fixed (i.e., capital is a state variable) and the controls (intermediate input and labor) are determined after the firm observes the productivity shock. Factor markets are competitive and factor prices are common across firms.
>
> **Assumption 2**. Firm *i* transforms inputs into output according to the following production technology[4] $Q_i = F(K_i, L_i, M_i, W_i) : R^4 \to R$. This function is twice continuously differentiable in L and M, and F_{LW}, F_{ML} and F_{MW} exist for all values $(K, L, M, W) \in R^4$.
>
> **Assumption 3**. Firm *i* takes the output prices for the homogeneous good as given.
>
> **Assumption 4**. The following inequality holds everywhere
>
> $$F_{LW}F_{ML} - F_{MW}F_{LL} > 0$$

The first assumption lays out the standard hypothesis on how inputs are optimally determined. Assumption 2 imposes standard regularity conditions on the production function. Assumption 3 essentially says that each firm is a price-taker and therefore faces a flat residual (inverse) demand curve - demand price elasticity is

[3] LP assume that output and input prices are common across firms. Therefore, these prices do not show up in (2).

[4] As in LP, I omit N_{it} from the production function technology for notational simplicity. This will not influence the results.

infinite. Finally, assumption 4 imposes an extra restriction on the production function derivatives.

Assumption 3 is certainly the most restrictive one, since many industries present some degree of product differentiation and market power. However, loosening assumption 3 is not such a simple task. In a perfectly competitive environment the monotonicity result is straightforward. With fixed output and input prices an increase in productivity implies higher output, which in turn implies higher demand for the intermediate input. In the context of imperfect competition this monotonicity property may no longer hold as firms realize that prices decrease as output goes up. Therefore, whether or not the intermediate input use increases as productivity goes up will also depend on the price sensitivity of consumers (i.e., demand elasticity). This is where this paper contribution comes in. I show that, under a certain form of imperfect competition, the monotonicity result is valid. To do so I impose the following assumptions:

Assumption 3'.

(a) Each firm produces a differentiated good.
(b) Each Firm faces a downward sloping demand curve for its differentiated product.
(c) A quantity change by one firm has a negligible effect on the price of any other firm.
(d) The price elasticity, define here as $-\sigma_i$ (where $\sigma_i > 1$), is constant along each firm's residual demand.

Assumption 3'(a) to 3'(c) describes the market configuration known as monopolistic competition (Chamberlin, 1933). In this market environment, firms are too small to influence market aggregates but still retain some market power due to product differentiation. Therefore each firm behaves as a monopoly on its downward sloping (residual) demand curve. Further, I assume the demand price elasticity is constant (see 3'(d)). Tirole (1988) shows that the constant price elasticity property can be obtained from a monopolistic competitive setup where a representative consumer has a CES utility function for the differentiated goods.[5]

[5] More specifically, Tirole makes the following assumptions: each monopolistic competitive firm produces one variety *i* and the representative consumer has the utility for the available varieties described as $U\left(\left(\sum_i Q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, Q_0\right)$, where Q_0 is an outside good. It is simple to show that the elasticity of substitution is constant and equal to σ. Then consumer optimization implies the following residual demand for each firm *i*: $Q_i = kP_i^{-\sigma}$, where *k* is a constant. Notice that demand price elasticity is constant and equal to – σ. Further, market clearing yields a formula for the price cost markup, which is given as σ/(σ-1). This setup gives an intuitively appealing interpretation. As the degree of product substitutability increases (i.e. σ increases) the demand function becomes flatter and firms lose market

Assumption 4'. The following inequality holds everywhere

$$F_{LW}F_{ML}-F_{MW}F_{LL}-\frac{1}{\sigma_i Q_i}(F_{LW}F_LF_M+F_{ML}F_LF_W-F_{MW}(F_L)^2-F_MF_WF_{LL})>0 \qquad (3)$$

Assumption 4' imposes an extra restriction on the production function and the demand elasticity. Note that the two assumptions above generalize assumption 3 and 4. Indeed, as we approach the competitive outcome, i.e. $\sigma\rightarrow\infty$, assumptions 3'and 4' become assumption 3 and 4. Now, I have the tools to show that the monotonicity condition works under imperfect competition.

Result. *If assumptions 1,2, 3'and 4' hold then the intermediate input demand function, $M_i(W_i,K_i)$, is strictly increasing in W_i.*

Proof. Firm i maximizes the following objective function with respect to L_i and M_i:

$$P_i(Q_i).Q_i-r_L.L_i-r_MM_i-r_KK_i$$

$P_i(Q_i)$ is the inverse demand for firm i's product, r_j is the return to factor j and $Q_i=F(K_i,L_i,M_i,W_i)$. Thus, dropping the subscripts indexing firms, one can write the first order conditions as

$$P(F(K,L(W),M(W),W)).F_j(K,L(W),M(W),W)=r_j\Big/\left(1-\frac{1}{\sigma}\right)\ ;j\text{=}L,M \qquad (4)$$

where F_j is the derivative of F with respect to factor j. Take the derivative of both sides of (4) with respect to W. Notice that the RHS of (4) is invariant to changes in W since r_j and σ are constant. Hence,

$$P\frac{dF_j}{dW}+F_j\frac{dP}{dF}\frac{dF}{dW}=0\ ;j\text{=}L,M. \qquad (5)$$

Next, divide this equation by P and derive the total derivatives. Thus,

$$\left(F_{jL}\frac{\partial L}{\partial W}+F_{jM}\frac{\partial M}{\partial W}+F_{jW}\right)+\frac{dP}{dF}\frac{F_j}{P}\left(F_L\frac{\partial L}{\partial W}+F_M\frac{\partial M}{\partial W}+F_W\right)=0\ ;j\text{=}L,M.$$

which can be written in terms of the demand elasticity as follows

power. In the extreme case, $\sigma\rightarrow\infty$, there is no product differentiation, firms become price-takers and the price cost ratio is one.

$$\left(F_{jL}\frac{\partial L}{\partial W}+F_{jM}\frac{\partial M}{\partial W}+F_{jW}\right)-\frac{1}{\sigma Q}F_j\left(F_L\frac{\partial L}{\partial W}+F_M\frac{\partial M}{\partial W}+F_W\right)=0\,;\ j{=}L,M. \qquad (6)$$

or, equivalently as

$$\begin{pmatrix} F_{LL}-\frac{1}{\sigma Q}(F_L)^2 & F_{LM}-\frac{1}{\sigma Q}F_LF_M \\ F_{ML}-\frac{1}{\sigma Q}F_MF_L & F_{MM}-\frac{1}{\sigma Q}(F_M)^2 \end{pmatrix}\begin{pmatrix}\frac{\partial L}{\partial W}\\ \frac{\partial M}{\partial W}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sigma Q}F_LF_W-F_{LW}\\ \frac{1}{\sigma Q}F_MF_W-F_{MW}\end{pmatrix}$$

Finally, from Cramer's rule

$$\frac{\partial M}{\partial W}=\det\begin{pmatrix} F_{LL}-\frac{1}{\sigma Q}(F_L)^2 & \frac{1}{\sigma Q}F_LF_W-F_{LW} \\ F_{ML}-\frac{1}{\sigma Q}F_MF_L & \frac{1}{\sigma Q}F_MF_W-F_{MW} \end{pmatrix}\Bigg/\det\begin{pmatrix} F_{LL}-\frac{1}{\sigma Q}(F_L)^2 & F_{LM}-\frac{1}{\sigma Q}F_LF_M \\ F_{ML}-\frac{1}{\sigma Q}F_MF_L & F_{MM}-\frac{1}{\sigma Q}(F_M)^2 \end{pmatrix}$$

The matrix in the denominator is the Hessian of the objective function. Thus, maximizing behavior implies that this matrix is negative semidefinite, i.e. the determinant of the Hessian is positive. From assumption 4', the numerator, which is given as

$F_{LW}F_{ML}-F_{MW}F_{LL}-\frac{1}{\sigma Q}(F_{LW}F_LF_M+F_{ML}F_LF_W-F_{MW}(F_L)^2-F_MF_WF_{LL})$, is strictly positive. Hence, $\partial M/\partial W$ is strictly positive and the monotonicity result follows■

Inequality (3) is simple (although tedious) to verify for any given production function. In fact, it is comforting to know that it holds, as I checked, for very familiar production functions as the CES and the Cobb-Douglas. It should be stressed that this paper's result does not alter the LP method. Rather, it broadens the set of assumptions under which the intermediate input proxy approach works.

3. An Example

In this section, I provide an example commonly found in the literature in order to highlight the problems caused by the assumption of perfect competition. This example also serves the purpose of illustrating the usefulness of this paper result. Measuring market power and productivity using plant-level data has been the object of many applied studies in industrial organization and trade (see the papers cited below).

In a seminal work, Hall (1990) shows that if capital is quasi-fixed and factor markets are competitive for the freely variable inputs then the log differentiated production function (1) can be written as

$$d\ln q_i = \gamma\, d\ln K_i + \mu(\alpha_{iL}(d\ln L_i - d\ln K_i) + \alpha_{iM}(d\ln M_i - d\ln K_i)) + e_i \qquad (7)$$

where e_i is an additively separable error term (productivity plus an i.i.d random variable) and α_{ij} is the cost share of input *j* relative to firm *i*'s total revenue. The specification above is consistent with the fact that capital is costly to adjust (i.e. capital is a state variable) and allows for the simultaneous estimation of returns to scale (γ), price-cost markup (μ), and productivity.

A number of researchers have used (7) to uncover the economically relevant parameters. Examples[6] include Harrison (1994) and Klette (1999). They differ in their solution to the simultaneity problem. The former author uses within estimator while the latter author uses an IV approach similar to the one developed by Blundell and Bond (2000). As mentioned in the introductory section these methodologies have a few shortcomings. Olley and Pakes offer a methodology that circumvents some of these problems but requires that zero-investment observations be removed from the data set. This requirement may significantly reduce the information available to the econometrician as zero-investment observations are a common feature in many data sets.

Hence, if one is not willing to lose a big chunk of the data set (and therefore reduce efficiency) and wants to avoid the standard approaches (within and IV estimators) inference problems, LP becomes the natural choice. However, notice that firm's optimizing behavior implies that the price-cost ratio is equal to $\sigma/(\sigma-1)$, where $-\sigma$ is the residual demand price elasticity. But in perfect competition this elasticity is infinite and, consequently, μ is one. Thus, findings of price-cost markups above one and the resulting productivity estimates are inconsistent with LP's assumption of perfect competition.

However, using the previous section's result it is possible to "fix" this consistency problem. Indeed, if one is willing to assume monopolistic competition and constant price elasticity, then assumption 3' is satisfied and, provided that assumptions 1, 2 and 4' are also satisfied, the monotonicity condition follows from the result derived in the previous section. Hence, the LP procedure (the inversion of the intermediate input demand function and the two-step algorithm) can be applied to estimate (7) without any inconsistencies.

[6] Another example that is closely related to the discussion in this section is Klette and Griliches (1996). They estimate equation (7) and assume monopolistic competition. However, as their focus was on the bias caused by unobserved price heterogeneity. They applied an IV approach to deal with the simultaneity bias.

4. Final Remarks

This paper provides a useful result for researchers interested in estimating production function parameters and markups using the LP framework. For, the monotonicity result developed here relaxes the perfect competition assumption imposed in the original LP paper. Indeed, this result guarantees consistency between a particular form of imperfect competition (monopolistic competition) and the LP procedure.

It is true that in some instances monopolistic competition may not be appropriate to model imperfect competition. However, the assumptions underlying monopolistic competition are certainly less restrictive than the ones underlying perfect competition. In this way, this paper can be viewed as a first step towards generalizing the monotonicity result. A natural extension of this work is to show that this result is robust to other forms of imperfect competition like Cournot and Bertrand.

Bibliography

Ackerberg, D., Caves K. and Frazer, G. (2005) "Structural Identification of Production Functions". *Working Paper,* University of Arizona.

Blundell, R. and Bond, S. (2000) "GMM Estimation with Persistent Panel Data: An Application to Production Functions" *Econometric Reviews*, **19**(3), 321-340.

Chamberlin, E. (1933). *The Theory of Monopolistic Competition*. Harvard university Press, Cambridge, MA.

Dunne, T. and Roberts, M. (1992) "Costs, Demand, and imperfect competition as determinants of plant-level output prices" Discussion Paper CES 92-5, Center for Economic Studies, U.S Bureau of the Census.

Griliches, Z. (1957) "Specification Bias in Estimates of Production Functions" *Journal of Farm Economics* **39**, 8-20.

Griliches, Z. and Mairesse, J. (1995) "Production Functions: the Search for Identification" NBER Working Paper No. 5067.

Hall, R. (1990) "Invariance Properties of Solow's Productivity Residual", in P. Diamond (ed.), *Growth, Productivity, Unemployment*, MIT Press, Cambridge, MA.

Harrison, A. (1994) "Productivity, Imperfect Competition and Trade Reform: Theory and Evidence" *Journal of International Economics* **36**, 53-73.

Klette, T.J. (1999) "Market Power, Scale Economies and Productivity: Estimates from a Panel of Establishment Data" *Journal of Industrial Economics*, **47**, 451-76.

Klette, T.J. and Griliches, Z. (1996) "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous" *Journal of Applied Econometrics* **11**(4), 343-61.

Levinsohn, J. (1993) "Testing the Imports-as-Market-Discipline Hypothesis" J*ournal of International Economics* **35**, 1-22.

Levinsohn, J. and Petrin, A. (2003) "Estimating Production Functions Using Inputs to Control for Unobservables" *Review of Economic Studies* **70**(2), 317-342.

Marschack, J. and Andrews, W.H. (1944) "Random Simultaneous Equations and the Theory of Production" *Econometrica* **12**(3,4), 143-205.

Melitz, M. J. (2000) "Estimating Firm-Level Productivity in Differentiated Product Industries" Working paper, Harvard University.

Olley, S. and Pakes, A. (1996) “The Dynamics of Productivity in the Telecommunications Equipment Industry” *Econometrica* **64**(6), 1263-1297.

Tirole, J.(1988). *The Theory of Industrial Organization*, MIT Press, Cambridge, MA.

Tybout, J. (2003) “Plant and Firm-level evidence on the ‘New’ Trade Theories”, in James Harrigan (ed.) *Handbook of International Trade*, Oxford: Basil-Blackwell.

Wooldridge J. (2005) “On Estimating Firm-Level Production Functions using Proxy Variables to Control for Unobservables”. Working Paper, Michigan State University.