Production functions with unobservable quantities: revisiting the omitted price bias.

Xulia González* and Daniel Miles-Touya†

Universidad de Vigo

Abstract

The standard practice in empirical industrial organization has been to approximate the firm's output as its revenue deflated by an industry price index. This approach introduces a measurement error to the extent that the industry prices do not accurately measures the firm's price level. Therefore, it might be of interest to test whether these deviations between the firm's price level and the industry price index affect the production function estimators. Some few attempts have been done in the literature with mixed results. In this paper we use a nonparametric test for assessing the relevance of the unobserved relative price measurment error. The test rejects the null hypothesis which suggests the measurement error is not conditionally independent of the production function explanatory variables, i.e. it may bias the estimators.

Keywords: production function productivity estimation
JEL Classification: L11,L60

*GRiEE and Universidade de Vigo. E-mail: xgzlez@uvigo.es.
†RGEA and Universidade de Vigo. E-mail: dmiles@uvigo.es.

INTRODUCTION

The estimation of production functions raises a set of empirical issues that have no straightforward solution (Ackerberg *et al.*, 2007; Syverson, 2011). One well-known issue concerns the measures of output and input, since firm level surveys typically do not report quantities nor prices. The standard solution has been to deflate the firm's revenues and intermediate goods expenditures by an industry price index, hence estimating a revenue function. However, this procedure may bias the coefficients of the production function to the extent that the difference between the firm's price and the industry price index may be correlated with de production function explanatory variables. More precisely, the revenue function estimators of the production function coefficients may be biased if the omitted price variable -the unobserved firm's relative price- were correlated with the firm's input choices (Klette and Griliches, 1996).

In most empirical applications the omitted price variable has been ignored or assumed away. Recently, the literature has focused on controlling this bias under the assumption that the omitted price variable is correlated with the input choices. The standard approach has been to assume a constant elasticity of substitution demand system which basically amounts to substituting the omitted price variable by the industry's aggregate demand in the empirical production function and rescaling the parameters by the demand elasticity (Klette and Griliches, 1996; hereafter KG; De Loecker, 2011). Alternatively, Grieco et al (2013) and Gandhi et al. (2013) exploit the firm's profit first order conditions to derive procedures to estimate the production function when prices are not observed.

Nevertheless, there is not to much empirical evidence assessing the correlation between the omitted price variable and the input choices, i.e. on how important it is in pratice. This issue may be of relevance since including additional structural assumptions to solve the omitted price variable may increase the risk of misspecification, hence deepening the bias problem.

In this empirical paper we assess the impact of the relative price measurement error on the production function estimators by testing whether the omitted price variable is correlated with the production function explanatory variables. We rely on the nonparametric test suggested by Henderson et al. (2008).

There have been some previous attempts in the literature which have tried to assess the impact of the relative price measurement error on the production function estimators. Naturally, this can only be studied when prices and quantities are available at the plant level, given that in this case it is straightforward to compare the production function and the revenue function estimates.

Foster et al. (2008) use information of physical ouptut reported by the North American Census of Manufactures to compare "quantity based" efficiency measures and revenue efficiency measures. Based on plant-level data on manufacturers of 11 homogeneous products they note that revenue-based and quantity-based productivity are highly correlated with each other and that these measures might suggest similar patterns with regard to industry dynamics (Syverson, 2011). Jaumandreu and Mairesse (2005) use a French and a Spanish Manufacture Panel Data Survey to compare the parameter estimates obtained when the firm's revenue is deflated by a firm output price index or an industry price index. Assuming that productivity can be proxied by a fixed effect, their results are in line with those suggested by Foster et al. (2008), in the sense that they do not find significant differences between the production function and revenue function parameter estimates. In other terms, these results are consistent with a model where neither the firm's optimal input decisions nor produtivity are correlate the the firm's relative price (the ratio between the firm price and the industry price index).

In contrast with these previous results, Ornaghi (2006) suggests that the unobserved relative price term may not be orthogonal to the input optimal decisions and/or productivity measures. Using the Spanish Manufacture Panel data and assuming that productivity is a fixed-effect in the production function, he compares the point estimates of the production function and revenue function when also considering the input relative price measurement error, i.e. intermediate goods expenditures are deflated by a common price index. His results suggest a statistically significant difference between these estimates.

In this paper we revisit this issue but following a different procedure to those of these

these previous papers. On the one side, we do not assume that the unobserved productivity can be specified as a fixed effect. Instead, we follow Olley and Pakes' (1996; OP hereafter) procedure and assume that productivity is a first order markov process. On the other side, our assessment is relies on a consistent nonparametric test built on the deviations between the firm's output price level and the industry price index. More precisely, the test is built on whether the residuals between the Olley and Pakes' (1996; OP hereafter) production function first stage estimator and the firm's revenue deflated by the industry price index are orthogonal to the prodution function explanatory variables. We apply this test to the Spanish Manufacture Panel Data survey and the test rejects the null hypothesis that the omited price variable is orthogonal to the input choices.

In order to implement the test we need a consistent estimate of the production function. However, the Spanish Manufacture panel data survey does not report the firm's output or input price level but its rate of change. As a consequence, as shown by Gonzalez and Miles (2013), it is not possible to recover a firm's specific price index to deflate revenues and recover quantities because we do not observe the base year price level (i.e., there is an initial condition problem). In the previous studies this issue has been overlooked and the initial condition problem has been addressed simply by fixing the unobserved base-year price level to a constant—for example, normalizing it to 1 (see e.g. Eslava *et al.*, 2004; Mairesse and Jaumandreu, 2005; Ornaghi, 2006; Dolado *et al.* 2012). However, such normalization implies that all firms in all industries have the same price level (i.e., there is neither intra- nor interindustry price dispersion) in the base year despite the existence of price heterogeneity in all other years. From an econometric standpoint, normalizing also implies assuming that the unobserved base-year price level is statistically independent of the production equation's right-hand-side variables –in other words, assuming that firms are positioned randomly in the base-year price distribution. This random hypothesis is rejected in Gonzalez and Miles (2013), a result that suggest treating the unobserved base-year price level as a fixed effect in the empirical production function equation (Gonzalez and Miles, 2013).

Therefore, the first stage of the procedure proposed by OP and ACF in order to account for the presence of the unobserved base-year price level fixed effect. The original OP/ACF procedure comprises two stages. In the first stage, a proxy function replaces the unobservable productivity in the production function, leading to a partially linear semiparametric model. In the presence of the base-year price level fixed effect the first stage of the OP procedure amounts to a partially linear semiparametric model with fixed effects (Su and Ullah, 2010). It is therefore necessary to remove the fixed effect in order to estimate consistently the nonparametric function that will be used in the second stage. Here we employ the fixed-effect profile, maximum likelihood estimator of Su and Ullah (2006; hereafter SU). The second stage of the OP procedure is not affected by the base-year price level fixed effect.

The paper is organized as follows: in the next section we discuss the difficulties that arise when deflating revenues and materials expenditures by an industry price index and also the initial condition problem attendant upon constructing a price index using rate-of-change data. In section three we present the estimation procedure and describe semiparametric fixed-effect methods. In section four we present the test. In section five we apply those methods to Spanish data and discuss the results before concluding in the paper's final section.

UNOBSERVED QUANTITIES: MEASUREMENT ERRORS

In this section we discuss the empirical problems that arises when firm level quantities are not observed. First, we discuss the problem that arises when the firm's nominal values are deflated by a common price index. Second, we present the problem that arises when we observe the firm's price rate of change but not the price levels.

Deflating by an industry price index

Following Klette and Griliches (1996), we assume that the production function for the ith manufacturing firm at time t can be represented by a Cobb-Douglas function:

$$Q_{it} = L_{it}^{\beta_l} K_{it}^{\beta_k} M_{it}^{\beta_m} \exp\left(\omega_{it} + \eta_{it}\right) \tag{1}$$

where Q_{it} denotes the quantity produced by firm i at time t, L_{it} the number of workers, K_{it} physical capital, and M_{it} quantities of materials. The parameters $\beta_l, \beta_k, \beta_m$ relate input choices to output; ω_{it} is the unobservable productivity shock (resulting from, e.g., managerial ability and entrepreneurship or expected downtime due to machine repairs); and η_{it} captures all those shocks that affect production but cannot be anticipated or predicted by the firm when making its input decisions–or are due to pure measurement error.

Expressed in logs, equation (1) can be written as

$$q_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \omega_{it} + \eta_{it} \tag{2}$$

where lowercase letters signify the logarithms of uppercase variables. The standard asumption in empirical industrial organization is that the production function is correctly specified, that is, $E\left(\eta_{it}|l_{it}, k_{it}, m_{it}, \omega_{it}\right) = 0$.

The main empirical shortcoming with regard to equation (2) is that quantities are seldom available in firm-level data sets, which instead report total revenue: $R_{it} = P_{it}Q_{it}$, where P_{it} is the ith firm's output price. Because firm data sets tend also to exclude firm price information, output is usually proxied as total revenue deflated by an industry price index P_{It}. That is, Q_{it} is measured as $\widetilde{R}_{it} = P_{it}Q_{it}/P_{It}$, which in logs is given by

$$\widetilde{r}_{it} = r_{it} - p_{It} = q_{it} + (p_{it} - p_{It})$$

or

$$\begin{aligned} q_{it} &= \widetilde{r}_{it} - (p_{it} - p_{It}) \\ &= \widetilde{r}_{it} - v_{it} \end{aligned}$$

Notice that this equation resembles a typical measurement error with respect to the dependent variable: the firm's logged output quantity at t is measured by the logged real revenue at t as corrected by the value of the relative price, $(p_{it} - p_{It})$.

Morever, let $E_{it} = M_{it}P_{it}^m$ be the firm's expenditure on materials and let P_{It}^m be the materials industry-wide price index for materials. Then the logged deflated expenditure is given by $e_{it} = m_{it} + (p_{it}^m - p_{It}^m)$, where the last term is an "omitted-input" price term in the empirical equation for the production function (Ornaghi, 2006).

Hence, if substitute the observable proxies in the production function equation, the result is

$$\widetilde{r}_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_m e_{it} + \omega_{it} + v_{it} + \eta_{it};$$

where v_{it} reflects the omitted price variables, $v_{it} = (p_{it} - p_{It}) - \beta_m \left(p_{it}^m - p_{It}^m\right)$.

As suggested by De Loecker (2011), in most empirical applications the omitted price variables have been ignored or assumed away. This implies arguing that the omitted price variables - the difference between the firm price and the industry price index- does not influence input decision nor is correlated with the firm's structural factors recovered through the productivity measure, i.e. the firm can not anticipate the difference between the firm's price and the industry price index when is deciding about its optimal input demands at t.

Therefore, it might be relevant to test whether the omitted price variables bias the production function parameter estimators before introducing additional structural assumptions. More precisely, we should test whether $E\left(v_{it}|l_{it}, k_{it}, e_{it}, \omega_{it}\right) = 0$.

Naturally, this test can only be performed if we can compare the estimates of the revenue function with those of the production function, i.e. when prices and quantities are available at the plant level. There are particularly few surveys that report quantities and prices, though some report the firm's price level rate of change. The firm's price level rate of change is usually used to obtain a firm price index in order to deflate revenues and obtain quantities. In the next subsection we discuss the problem that arises when we observe the firm's level price rate of change but not the firm's price level.

Price rate of change: the unobserved base-year price level

Some firm-level data sets—such as the Spanish ESEE, the Bank of Italy's Survey on Italian Manufacturing Firms (INVIND), the Colombian Encuesta Anual de Manufacturas (EAM), and the French manufacturing survey—report the rate of change in the firm's output price. Here we discuss the initial condition problem that arises when the rate of price change is used to derive a firm-specific price index, which is used to deflate nominal values.

Following Eslava *et al.* (2004), among others, we let the price information reported by a firm-level panel dataset be the rate of change, $\triangle P_{ijt}/P_{ijt-1}$, here $\triangle P_{ijt} = P_{ijt} - P_{ijt-1}$, j indexes the number of products produced by firm i and P_{ijt} is the price of product j charged by firm i at time t. The rate of change in the firm's output price can be obtained by using a Tornqvist index, which is a weighted average of the growth rate of individual products:

$$\frac{\triangle P_{it}}{P_{it-1}} = \sum_{j=1} s_{ijt} \left(\triangle P_{ijt}/P_{ijt-1}\right), \qquad t = 1, ..., T;$$

here s_{ijt} is any weighting function (e.g., the revenue or market share of product j of firm i at time t) that relates the product j to the firm's total.

The firm's output price level for each period is obtained from the recursion formula

$$P_{it} = (1 + \triangle P_{it}/P_{it-1})P_{it-1}, \qquad t = 1, ..., T,$$

where $\triangle P_{it}/P_{it-1}$ is given in (4) and backward induction yields the output price level at time t

$$P_{it} = \prod_{t=1}^{t} \left(1 + \triangle P_{it}/P_{it-1}\right) P_{i0}.$$

Here P_{i0} is the base-period price level and the firm's specific price index is now given by P_{it}/P_{i0}. Hence the output measure is equal to the firm's revenue, $R_{it} = P_{it}Q_{it}$, deflated by the firm's specific price index:

$$R^*_{it} = \left(P_{it}Q_{it}\right)/\left(P_{it}/P_{i0}\right) = P_{i0}Q_{it};$$

in logs, we have

$$q_{it} = r^*_{it} - p_{i0}. \tag{3}$$

It should be clear that the firm's logged output quantity at t is measured by the logged real revenue at t as corrected by the value of the base-year price level. The empirical limitation of this approach is that price levels are usually not observed in surveys that report the rate of change in prices; that is, p_{i0} is unavailable. Yet this issue has been overlooked by previous studies, which simply set the base-year price level to a particular constant—for instance $P_{i0} = 1$ for all i at $t = 0$, which implies that $p_{i0} = 0$.

Equation (3) , however, embodies a typical measurement error with respect to the dependent variable. The standard approach to correct for such an error is to assume that it is statistically independent of the explanatory variables and hence can be disregarded (Wooldridge, 2010). Note that this assumption is implicitly invoked when the base-year price level is set to any given constant.

Furthermore, the materials price index suffers an analogous measurement error problem: the base-year price level of materials is not observed in surveys that report data on the rate of change for prices. Thus we have $m_{it} = e_{it} - p_{i0}^{m}$, where e_{it} is the log of (deflated) expenditures on materials and p_{i0}^{m} is the log of the base-year price level of materials. Because this measurement error concerns an explanatory variable, it may well bias the parameter estimators.

Here we treat the unobservable base-year price level of output and materials as measurement error fixed effects. Hence, if output is measured by $q_{it} = r_{it}^{*} - p_{i0}$ and materials by $m_{it} = e_{it} - p_{i0}^{m}$, substituting in equation (2)

$$r_{it}^{*} = \beta_l l_{it} + \beta_k k_{it} + \beta_m e_{it} + \omega_{it} + \tau_i + \eta_{it},$$

where $\tau_i = p_{i0} - \beta_m p_{i0}^{m}$ captures the base-year price levels of materials and output.

ESTIMATING THE PRODUCTION FUNCTION WITH PARTIALLY OBSERVED FIRM-LEVEL PRICES

In order to perform the test we first need a consistent to estimate the production function. Here we rely on semiparametric frist stage of the approach proposed by Olley and Pakes (1996) under the timing assumptions suggested by Ackerberg *et al.* (2006). In this section we briefly describe the original OP procedure (with the ACF assumptions) and introduce the modifications needed when the base-year price level is considered a fixed effect.

The original OP approach considers firms that make production choices to maximize the present discounted value of current and future profits. A firm's production function resembles equation (1), where the unobserved productivity ω_{it} is assumed to follow an exogenous first-order Markov process, i.e. $\Pr\left(\omega_{it} | \left\{\omega_{it-j}\right\}_{j=1}^{t}\right) = \Pr\left(\omega_{it} | \omega_{it-1}\right)$, which is stochastically increasing in ω_{it-1}. Inputs to be used in period t can be classify in flexible inputs, which can be adjusted in every period, e.g. materials, or quasi-‡exible, subject to adjustment frictions, which are chosen at $t-1$, e.g. capital or labor. Furthermore, inputs can either be dynamic, in which case the current period's input choices affect the firm's future profits (these are so-called state variables), or nondynamic. Finally, firms are assumed to operate in perfectly competitive output and input markets. In this economic environment, the firm's profit maximization problem results in an investment policy rule that depends on the firm's unobserved productivity and the state variables—namely, $i_{it} = f_t(\omega_{it}, k_{it}, l_{it})$ (Ackerbert *et al.*, 2007).

The OP model relies crucially on the notion that this investment policy rule can be inverted to proxy for unobserved productivity in the production function equation via a function of investment and the state variables: $\omega_{it} = f_t^{-1}(i_{it}, k_{it}, l_{it})$. Pakes (1994) gives conditions for invertibility when only one unobservable affects firm behavior (i.e., a scalar unobservability assumption); the implication is that the investment function is strictly increasing in ω_{it} in the region where i_{it} is positive.[1]

Under these assumptions, the original OP method consists of two stages. In the first stage, the unobservable productivity in the production function is replaced by the inverse of the investment function (i.e., the proxy function) to yield a partial linear model. That is, in light of (2) we have

$$q_{it} = \beta_m m_{it} + \phi_t(i_{it}, k_{it}, l_{it}) + \eta_{it}$$

where $\phi_t(i_{it}, k_{it}, l_{it}) = \beta_l l_{it} + \beta_k k_{it} + f_t^{-1}(i_{it}, k_{it}, l_{it})$. In principle, this first stage could be used to estimate a semiparametric version of the production function, i.e. we can obtain the coefficients for the nondynamic variable inputs—that is, the intermediate inputs coefficient β_m as well as ϕ_t.[2].

[1] Recent papers have relaxed this scalar unobservable assumption by admitting additional unobservables in the investment equation (DeLoecker, 2007; Ackerberg et al., 2007; Huang and Hu, 2011; Aguirregabiria and Alonso-Borrego, 2013).

[2] The OP method's second stage identifies the parameters of the dynamic inputs by relying on the moment conditions that result from (a) the first-order Markov assumption, $\omega_{it} = E(\omega_{it}|\omega_{it-1}) + \varepsilon_{it} = g(\omega_{it-1}) + \varepsilon_{it}$

Ackerberg *et al.* (2006) discuss the ability of the OP procedure to identify the variable input coefficients in the first stage. Under OP assumptions, the variable input demand could be a function of productivity, in which case—conditional on the first stage's proxy function—no variability would be left for intermediate materials. That is, $m_{it} = g_t(\omega_{it}, k_{it}, l_{it})$ so $m_{it} = g_t\left(f_t^{-1}(i_{it}, k_{it}, l_{it}), k_{it}, l_{it}\right) = z_t(i_{it}, k_{it}, l_{it})$, which implies that β_m is not identified in the first stage because there remains no variation conditional on the nonparametric function.

In order to allow for identification in the first stage, ACF propose additional timing assumptions concerning when productivity shocks hit the firm and when inputs are chosen.[3] The idea is to introduce an independent source of cross-sectional variability by assuming that input decisions are based on an information set other than that of investment. So, suppose that ω_{it} evolves between subperiods $t-1$,$t-b$, and t according to a first-order Markov process and tha m_{it} is chosen at $t-b$. In this case the firm's optimal material input will be a function not of ω_{it} but rather of ω_{it-b}, which resolves the multicollinearity problem and allows the first stage to be identified. The intuition is that intermediate goods are chosen without perfect information about ω_{it}, and this incomplete information is what moves m_{it} independently of the nonparametric function. Hereafter, we incorporate these

and (b) the time to build assumption, $E(k_{it}\varepsilon_{it}) = E(l_{it-1}\varepsilon_{it}) = 0$. We can recover the residual ε_{it} by nonparametrically regressing $\omega_{it}(\beta_k, \beta_l)$ on $g(\omega_{it-1}(\beta_k, \beta_l))$, given that

$$\omega_{it}(\beta_k, \beta_l) = \widehat{\phi}_t - (\beta_l l_{it} + \beta_k k_{it})$$

and then estimate the coefficients by minimizing the sample analogue of the moment conditions,

$$\frac{1}{N}\frac{1}{T}\sum \varepsilon_{it}(\beta_k, \beta_l)\begin{bmatrix} k_{it} \\ l_{it-1} \end{bmatrix}$$

using standard GMM techniques.

[3] ACF present an alternative approach to dealing with the problem of multicollinearity: estimate only the nonparametric function in the first stage and then estimate all the parameters in the second stage. Thus the first-stage estimation would be given by

$$q_{it} = \phi^*_{it}(i_{it}, k_{it}, l_{it}, m_{it}) + \eta_{it}$$

and the second stage would use moment conditions to estimate the coefficients.

ACF assumptions into the OP procedure.[4]

Moreover, in OP original setting, all firms are confronted to the same prices because they operate in a perfectly competitive markets. As a consequence, the firm price level does not constitute a state variable in the investment function. In our context, there is price heterogeneity which can imply that prices should be considered as a state variable. However, we follow De Loecker (2011) and the price structural underpinings given by Foster et al. (2008) and Roberts and Supina (2000) to argue that the firm prices are not state variables in the investment optimal demand function, i.e. prices are conditionally serially uncorrelated.

De Loecker (2011) assume that the unobserved firm level prices are picked up by the variation in inputs and by aggregate demand, which he argues are not conditionally serially correlated, i.e. aggregate demand is not a state variable. In our case, the aggregate demands in the industry-year intersection are constant for the firms in the same industry and, therefore aggregate demand is capture through year dummies[5]. Moreover, Forster's et al. (2008) explicitly derive the firm price as a function of productivity and demand shifters, which suggest that, conditioning on these variables, the price idyosincratic shock is conditionally serially uncorrelated. Additionally, Roberts and Supina (2000) present a similar price equation as Foster at al. (2008) and suggest that the firm's price indyosincratic shock may not serially correlated once conditioning on firm's structural characteristics that change slowly over time, e.g. productivity, which embodies the producers idyosincratic tecnologies that explain prices movements.

As a consequence, the production function with fixed-effects measurement error, τ_i is given by

$$r^*_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_m e_{it} + \omega_{it} + \tau_i + \eta_{it}.$$

[4] The following ACF assumptions support their conclusions: the realization of materials cost shocks are not conditionally serially correleted and are realized after the investment decision takes place; and capital and labor markets are perfectly competitive.

[5] Notice that this is equivalent to assume that firms are confronted to a constant substitution demand function (Klette and Griliches, 1996; De Loecker, 2011).

and $i_{it} = f_t(\omega_{it}, k_{it}, l_{it}, D_{it})$, where D_{it} captures observable demand shifters and time dummy variables.[6]

The foregoing discussion indicates that the first-stage equation can be written as

$$\begin{aligned} r^*_{it} &= \beta_l l_{it} + \beta_k k_{it} + \beta_m e_{it} + f_t^{-1}(i_{it}, k_{it}, l_{it}, D_{it}) + \tau_i + \eta_{it} \\ &= \beta_m e_{it} + \phi_t(i_{it}, k_{it}, l_{it}, D_{it}) + \tau_i + \eta_{it}. \end{aligned}$$

Here $\phi_t(i_{it}, k_{it}, l_{it}, D_{it}) = \beta_l l_{it} + \beta_k k_{it} + \omega_{it}$ and $\omega_{it} = f_t^{-1}(i_{it}, k_{it}, l_{it}, D_{it})$.[7]

The presence of a base-year price level fixed effect in the production function affects the estimation procedure. We next show how to modify the first stage of the OP procedure by means of semiparametric, partially linear, panel data fixed-effect models that enable us to estimate not only the first-stage nonparametric function but also parameters of the nondynamic inputs. The nonparametric function is then used in the second stage to estimate the dynamic inputs, just as in the original OP/ACF procedure.

Partially linear panel data models

The first stage of the OP method can be described as a partially linear panel data fixed-effect model. We employ the notation most commonly used in semiparametric partial linear

[6] We remark that measurement error in the materials input may affect the derivation in Levinsohn and Petrin (2003). Their procedure derives the proxy function from the optimal materials demand decision, $m_{it} = d_t(k_{it}, l_{it}, \omega_{it})$ so $\omega_{it} = d_t^{-1}(k_{it}, l_{it}, m_{it})$. If there is a base-year measurement error, then $\omega_{it} = d_t^{-1}(k_{it}, l_{it}, e_{it} - p^m_{io}) = d_t^{-1}(k_{it}, l_{it}, e_{it}, \tau_i)$; this expression characterizes a partially separable panel data model.

[7] Observe that if the unobservable base-year price level is considered a state variable, then we have a partially separable nonparametric panel data model. That is, if $i_{it} = f_t(\omega_{it}, k_{it}, l_{it}, \tau_i)$ so $\omega_{it} = f^{-1}(k_{it}, l_{it}, i_{it}, \tau_i)$, then subsituting into the production function equation yields $r^*_{it} = \phi(i_{it}, k_{it}, l_{it}, m_{it}, \tau_i) + \eta_{it}$ under the ACF alternative specification for the OP's procedure first stage. Unfortunately, existing methods for nonseparable models recover only the average derivatives' estimators. In ongoing research, Evdokimov (2010), has proposed a method for estimating the nonparametric function ϕ.

models:[8]

$$y_{it} = x_{it}'\beta + m\left(z_{it}\right) + \tau_i + \eta_{it}, i = 1, ..., n; t = 1, ..., T_i.$$

Here y_{it} represents r_{it}^*; x_{it} is a $p \times 1$ vector of explanatory variables, including materials inputs e_{it}; $m\left(\cdot\right) = \phi\left(\cdot\right)$; and $z_{it} = \left(i_{it}, k_{it}, l_{it}, D_{it}\right)$, which is a $q \times 1$ vector.[9]

Recent years have seen an increasing number of papers that propose alternative estimation methods for use with fixed-effect nonparametric or semiparametric panel data models (Su and Ullah, 2010). Each of these papers proposes an alternative method of estimating the linear or nonparametric component. For example, Li and Stengos (1996) use a kernel instrumental variable to estimate the linear—but not the nonparametric—component. Henderson *et al.* (2008) propose a similar iterative kernel estimator for the nonparametric function that is, however, computationally more demanding.

In this paper we apply two straightforward methods for estimating the unknown component of the partial linear model. One is based on profile likelihood estimation for the nonparametric component (Su and Ullah, 2006); the other is based on series approximation (Baltagi and Li, 2002). We have chosen these two methods because they are computationally less costly and have good asymptotic properties (Su and Ullah, 2010; Gao, 2012).

Su and Ullah (2006) propose estimating the semiparametric fixed-effect model by means of profile maximum likelihood (i.e., profile least squares). The local linear approximation of their equation can be written as

$$\begin{aligned} y_{it} &= x_{it}'\beta + m\left(z\right) + \frac{\partial m}{\partial z}\left(z_{it} - z\right) + \tau_i + \eta_{it} \\ &= x_{it}'\beta + Z_{it}\left(z\right)\delta\left(z\right) + \tau_i + \eta_{it} \end{aligned}$$

where $\frac{\partial m}{\partial z} = \left(\frac{\partial m}{\partial z_1}, ..., \frac{\partial m}{\partial z_q}\right)$, $Z_{it}\left(z\right) = \left[1, \left(z_{it} - z\right)'\right]_{1\times(q+1)}$ and $\delta\left(z\right) = \left(m\left(z\right), \frac{\partial m}{\partial z}\right)'$.

[8] These methods were developed for data from balanced panels: we have extended these procedures to unbalanced panels.

[9] Although OP, ACF, and Levinsohn and Petrin (2003) all assume that the functional form of the investment proxy is time dependent, in their empirical applications this form is presumed not to be time dependent. In fact, it is possible (as suggested by a referee) for these functions to depend on t, but that would require an estimation methodology not yet available in the literature.

The idea behind this approach is to profile out the individual effect and the linear parameters, $\theta = (\tau', \beta')'$ and then consider the concentrated least squares for $\delta(z)$. Therefore, if θ is known the estimate of $\delta(z)$ is given by

$$\delta_\theta(z) = \arg\min_{\delta \in R^{q+1}} \sum_i \sum_t \left(y_{it} - \tau_i - x'_{it}\beta - Z_{it}(z)\,\delta(z)\right)^2 K_H(z_{it}, z),$$

where $K_H(z_{it}, z) = \prod_{k=1}^{q} h_k^{-1} k\left(\frac{z_{it,k}-z}{h_k}\right)$for $k(\cdot)$ a univariate kernel function and subject to the identification condition $\sum_i \tau_i = 0$. Thus,

$$\delta_\theta(z) = \left[\sum_i \sum_t Z_{it}(z)' K_H(z_{it}, z)\, Z_{it}(z)\right]^{-1} \sum_i \sum_t Z_{it}(z)' K_H(z_{it}, z)\left(y_{it} - \tau_i - x'_{it}\beta\right).$$

However, $\delta_\theta(z)$ is not an operational term because it depends on the unknown parameter θ. Hence this method profiles out $m(z)$ by estimating θ as

$$\widehat{\theta} = \arg\min_\theta \sum_i \sum_t \left(y_{it} - \tau_i - x'_{it}\beta - m_\theta(z_{it})\right)^2$$

where $m_\theta(z_{it})$ is the first component of $\delta_\theta(z)$. Given the profile estimates $\widehat{\theta} = \left(\widehat{\tau}', \widehat{\beta}'\right)'$, the profile likelihood estimator of $m(z)$ is

$$\begin{aligned}
\widehat{m}(z) &= m_{\widehat{\theta}}(z) = e'\delta_{\widehat{\theta}}(z) \\
&= e'S(z)\left(\mathbf{y} - D\widehat{\tau} - \mathbf{X}\widehat{\beta}\right)
\end{aligned}$$

in matrix form, where $e = (1, \mathbf{0}_{q\times 1})_{(q+1)\times 1}$,

$$\begin{aligned}
S(z) &= \left[Z(z)' K_H(z)\, Z(z)\right]^{-1} Z(z)' K_H(z) \\
Z(z) &= \left(Z_{11}(z)' \ldots Z_{1T_1}(z)' \ldots Z_{nT_n}(z)'\right)'_{R\times(q+1)}, \\
K_H(z) &= diag\left[K(z_{11} - z), K(z_{12} - z) \ldots K(z_{1T_1} - z) \ldots K(z_{nT_n} - z)\right]_{R\times R} \\
D_{R\times n-1} &= [(-\iota_{n-1}, I_{n-1}) \otimes \iota_{T_i}]_{i=1,\ldots,n} \\
\mathbf{y} &= (y_{11,\ldots} y_{nT_n})'_{R\times 1}
\end{aligned}$$

for $R = \sum_{i=1}^{n} T_i, , \tau = (\tau_2, \ldots \tau_n)'$ and $\mathbf{X}_{R\times p}$.

The parameter profile estimates are given by

$$\begin{aligned}\widehat{\beta} &= \left[\widetilde{X}'\widetilde{X}\right]^{-1}\widetilde{X}'\widetilde{\mathbf{y}} \\ \widehat{\tau} &= \left[D^{*\prime}D^{*}\right]^{-1}D^{*\prime}\left(\mathbf{y}^{*}-X^{*\prime}\widehat{\beta}\right).\end{aligned}$$

Here the asterisk * indicates the residual after regressing the respective variable on the matrix defined by $S = [s(z_{it})] = [e'S(z_{it})]_{\ i=1,\ldots,n;t=1,\ldots.T_n}$ and the tilde $\widetilde{\cdot}$ denotes the residuals of regressing X^{*} and $\mathbf{y}^{*}$ on D^{*}. Su and Ullah (2006) demonstrate the consistency of the nonparametric estimator and +the $\sqrt{n}-$consistency of the parameters estimator. In Su and Ullah (2007) these authors develop a similar estimation procedure for random-effects models.

A TEST FOR THE OUTPUT RELATIVE PRICE TERM

In this section we built the test to assess the relevance of the relative price measurement error. The test relies in evaluating the residuals defined as the difference between the firm's revenue deflated by an industry index and the consistent estimate of the production function.

Let the production function be stated as in the first stage of the OP procedure be given by

$$\begin{aligned}q_{it} &= \beta_m m_{it} + \phi_t(i_{it}, k_{it}, l_{it}, D_{it}) + \eta_{it} \\ &= \phi_t^{*}(z_{it}) + \eta_{it}\end{aligned}$$

where $E(q_{it}|z_{it}) = \beta_m m_{it} + \phi_t(z_{it}) = \phi_t^{*}(z_{it})$, $z_{it} = (i_{it}, k_{it}, l_{it}, D_{it}, m_{it})$ and $\phi_t(\cdot)$ is an unknown functions. In the previous section we described how we can recover a consistent estimate $\phi_t^{*}(\cdot)$ when we only observe the firm's level price rate of change.

The firm's output as measured by the firm's revenue deflated by the industry price index is given by

$$q_{it} = \widetilde{r}_{it} - v_{it},$$

and substituting in the first stage of the OP procedure we have

$$\widetilde{r}_{it} = \phi_t^* (z_{it}) + v_{it} + \eta_{it}.$$

Therefore, if $E(v_{it}|z_{it}) = 0$ then case $E(q_{it}|z_{it}) = E(\widetilde{r}_{it}|z_{it})$. As a consequence, the test can be based in a statistic derived from the orthogonality condition,

$$E(\widetilde{r}_{it} - \phi_t^* (z_{it}) | z_{it}) = 0,$$

where all its components are observable or can be consistently estimated.

In order to derive the test statistic we follow Henderson et al. (2008) see also Li and Sun, 2013). Let $u_{it} = (\widetilde{r}_{it} - \phi^* (z_{it}))$ be the residual between the firm's revenue deflated by the industry price index and the OP first stage production function. Then, we want to test

$H_0 : E(u_{it}|z_{it}) = 0$

$H_A : E(u_{it}|z_{it}) \neq 0$

The proposed test is based on the sample analogue of $J = E\{u_{it}E(u_{it}|z_{it}) f(z_{it})\} = E\left\{[E(u_{it}|z_{it})]^2 f(z_{it})\right\}$. If the model is assumed to be correctly specified, then the statistic equals zero under H_0 -no omitted variable bias- but exceeds zero under H_A.

In order to implement the statistic, let $\widehat{\phi}^* (z_{it}) = \widehat{\beta}_m m_{it} + \widehat{\phi}(i_{it}, k_{it}, l_{it}, m_{it})$ be consistent estimators of $\phi^* = \beta_m m_{it} + \phi(i_{it}, k_{it}, l_{it}, m_{it})$ described in the previous section. Then, a consistent estimator of u_{it} is given by $\widehat{u}_{it} = \widetilde{r}_{it} - \widehat{\phi}^* (z_{it})$ where $\widetilde{r}_{it}$ is the firm's revenue deflated by the industry price index.

A feasible test statistic is given by

$$\widehat{J} = \frac{1}{R} \sum_{i=1}^{n} \sum_{t=1}^{T_i} \widehat{u}_{it} E_{-it} (\widehat{u}_{it}|z_{it}) \widehat{f}_{-it} (z_{it})$$

where $R = \sum_{i=1}^{N} T_i$ and where a wild bootstrap procedure is used to approximate the finite sample null distribution of $\widehat{J}$ (Li and Sun, 2013; Henderson et al., 2008).

The null hypothesis implies that the relative price omitted term does not bias the production function parameter estimators, i.e. the production function is consistently estimated by the revenue funciton. Therefore, the bootstrap is based on the revenue production function residuals. Let $\widehat{\mathbf{u}}^r = \left(\widehat{u}_{11}^r, ..., \widehat{u}_{nT_n}^r\right)'_{R\times 1}$ be the residuals of the revenue function as in the first stage of the OP procedure,

$$\widetilde{r}_{it} = \widehat{\beta}_m^r m_{it} + \widehat{\phi}^r \left(i_{it}, k_{it}, l_{it}\right) + \widehat{u}_{it}^r$$

from where we define the bootstrap residuals based on a two point wild bootstrap

$$u_{it}^{r*} = \begin{cases} \left(\left(1-\sqrt{5}\right)/2\right)\widehat{u}_{it}^r & p = \left(1+\sqrt{5}\right)/\left(2\sqrt{5}\right) \\ \left(\left(1+\sqrt{5}\right)/2\right)\widehat{u}_{it}^r & 1-p \end{cases}$$

We generate the bootstrap sample $\{\widetilde{r}_{it}^*, z_{it}\}_{i=1,..,n;t=1,...T_i}$, as

$$\widetilde{r}_{it}^* = \widehat{\beta}_m^r m_{it} + \widehat{\phi}^r \left(i_{it}, k_{it}, l_{it}, m_{it}\right) + u_{it}^{r*}$$

where $\widehat{\beta}^r, \widehat{\phi}^r$ are the revenue function estimates in the original sample. This bootstrap sample $\{\widetilde{r}_{it}^*, z_{it}\}_{i=1,..,n;t=1,...T_i}$ is then used to estimate the revenue function bootstrap residuals, $\widehat{u}_{it}^{r*} = \widetilde{r}_{it}^* - \widehat{\beta}_m^{*r} m_{it} - \widehat{\phi}^{*r}\left(i_{it}, k_{it}, l_{it}, m_{it}\right)$, where $\left(\widehat{\beta}^{*r}, \widehat{\phi}^{*r}\right)$ are the bootstrap revenue function estimates. The bootstrap test statistic $\widehat{J}^*$ is obtained just as is $\widehat{J}$ except that the residual, $\widehat{u}_{it}$, is replaced by the revenue function bootstrap residual $\widehat{u}_{it}^{r*}$. This process is repeated a large number of times after which the empirical bootstrap distribution is then used to approximate the null distribution of the test statistic $\widehat{J}$. For each bootstrap sample we perform the test and repeat for B bootstraps. The bootstrap p-value is given as $B^{-1}\sum I\left(\widehat{J}^* > \widehat{J}\right)$.

In the next section we apply the methods just described to data from a survey of Spanish manufacturing firms.

DATA AND RESULTS

Our data comes from the Encuesta Sobre Estrategias Empresariales (ESEE) 1991–2006 survey, an unbalanced firm-level panel of Spanish manufacturing firms that is sponsored by the Ministry of Industry. These data have been frequently used to estimate the production function. We selected a subsample consisting of small and medium-size firms (i.e., firms with 10–199 workers) that exhibited positive investment levels and at least two consecutive periods in the sample. This selection criteria yielded a total of 10,484 observations of 1,616 firms belonging to ten different industrial sectors.

In Figure 1we compare 75th and 25th quantiles firm's price index obtained normalizing the base year price level, i.e. $P_{0i} = 1$ for all i, and the Industry price index. As it can be observed from the Figure, the normalization restricts all prices to be the same in the base year though there are singificant price differences in subsequent years. These graphs could be suggesting two things: first, that the production function parameters estimators defined using the revenue function may be bias; second, that the normalization of the unobserved base-year price level to a particular constant for all firms in all industries may affect the production function parameter estimators.

[Insert Figure 1 about here]

In Table 1we present the results of implementing the test to assess whether the unobserved relative price term bias the production function parameter estimators under three different bandwith parameters..

[Insert Table 1 about here]

Overall, the test rejects the null hypothesis that the unobserved relative price term is uncorrelated with the production function explanatory variables. In other terms, recovering the production function parameters from the revenue function may lead to uncorrect results. Another interpretation of the Table 1 results is that the conventional functional forms adopted in the empirical implementation of the OP/ACF procedure are incorrect (Bierens, 2009). That is, because the nonparametric test can detect overall misspecification of a functional form, the test could imply either that the unobserved relative price term significantly affects the parameter estimators (assuming a correct functional form specification) or that the overall functional forms are not correct.

CONCLUSION

A limitation of the empirical research on production functions is that output and input quantities are not observed. When firm-level price information is unavailable, the standard approach is to measure the firm's output as revenue divided by an industry-level deflator (revenue production function). However, this approach introduces an omitted variable problem defined by the difference between the firm's price and the industry price, i.e. an unobserved relative price term. There have been some few attempts in the literature trying to assess whether this omitted variable term biases the production function parameter estimators. These papers used a fixed-effect productivity specification and assessed the statistical relevance of the unobserved price term by comparing the estimates of the production function and the revenue function, reporting mixed results.

In this paper we use a formal nonparametric test to assess whether the unobserved relative price term may bias the production function parameter estimators. Using a panel data from the Spanish manufacturing industry we reject the null hypothesis that states that the unobserved relative price term is uncorrelated with the production function explanatory viariables. In other terms, the revenue function parameter estimators report bias production function parameters.

APPENDIX A: SURVEY AND DEFINITION OF VARIABLES

The Encuesta Sobre Estrategias Empresariales (ESSE) 1990–2006 is a panel of firms. The raw data set consists of 4,357 manufacturing firms and a total of 30,827 observations. At the beginning of this survey in 1990, 5% of firms with fewer than 200 workers were sampled randomly by industry and size strata. All such firms were asked to participate, which they did at a rate of about 70%. The initial sample properties have been maintained in subsequent years because exit attrition is balanced by replacing exiting firms with newly created firms that satisfy the initial sampling criteria as in the first year.

For the research reported in this paper, we selected a subsample consisting of small and medium-size firms that are present in the panel data for at least two consecutive years; this subsample includes 15,757 observations for 2,616 firms. Because the OP procedure considers only firms with positive investment, we eliminated from the sample all firms that did not exhibit positive investment. Hence our final sample contained 10,484 observations for 1,616 firms.

The variables are defined as follows.

- *Capital.* Capital at current replacement values K_{it} is computed recursively from an initial estimate and using data on current investments in equipment goods I_{it}. We update the value of the past stock of capital by means of the price index of investment in equipment goods p_{It} as $K_{it} = (1-\delta)(p_{It}/p_{It-1})K_{it-1}+I_{it-1}$, where δ is an industry-specific estimate of the rate of depreciation. Capital in real terms is obtained by deflating capital at current replacement values by the price index of investment in equipment goods.

- *Investment.* The value of current investments in operative capital includes equipment goods but excludes buildings, land, and financial assets. The magnitude is deflated by the price index of investment (the equipment goods component of the index of industry prices computed and published by the Spanish Statistic Institute, INE).

- *Market dynamism.* Firms are asked to assess the current and future situation (slump, stability, or expansion) of up to five separate markets in which they operate. The market dynamism index is computed as a weighted average of the responses.

- *Materials.* Value of intermediate consumption (including raw materials, components, energy, and services) deflated by a firm-specific price index of materials.

- *Output.* Value of produced goods and services, computed as sales plus the variation of inventories and deflated by a firm-specific price index of output.

- *Employment.* Number of full-time plus half of part-time workers as of December 31.

- *Demand shifters*: firms are asked to assess the current and future situation of up to 5 separate markets which they operate: contraction, stability or expansion.

- *Materials price index*: Firm-specific price index for intermediate consumption: firms are asked about the price changes that occurred during the year for raw materials, components, energy, and services. The price index is computed as a Paasche-type index.

- *Output price index*: Firm-specific price index for output. Firms are asked about the price changes they made during the year in up to 5 separate markets in which they operate. The price index is computed as a Paasche-type index.

REFERENCES

Ackerberg, D., K. Caves and G. Frazer (2006), "Structural Identification of Production Functions" *mimeo.*

Ackerberg, D., C. L. Benkard, S. Berry, and A. Pakes (2007), "Econometric Tools for Analyzing Market Outcomes" in James Heckman and Edward Leamer, eds., Handbook of Econometrics, Vol. 6(1), Amsterdam: North-Holland, 4171–4276.

Aguirregabiria, V. (2009), "Econometric issues and methods in the estimation of production functions" MPRA Paper 15973.

Aguirregabiria, V. and Alonso-Borrego, C. (2013), "Labor contracts and flexibility: evidence from a labour market reform in Spain" *Economic Inquire*, forthcoming.

Bierens, H. (2009), "Consistent specification tests" *mimeo.*

De Loecker, J. (2007), "Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity" *NBER* WP 13155.

De Loecker, J. (2011), "Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity" *Econometrica* 79(5), 1407–1451.

Dolado, J. Ortigueira, S. and Stucchi (2012), "Does dual employment protection affect TFP? Evidence from Spanish manufacturing firms" *CEPR Discussion Papers,* 8763.

Doraszelski, U. and J. Jaumandreu (2013), "R&D and Productivity: Estimating Endogenous Productivity" *The Review of Economic Studies,* forthcoming.

Eslava, M., Haltiwanger, J., Kugler, A. and Kugler, M. (2004), "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia" *Journal of Development Economics* 75 (2), 333–372.

Evdokimov, K. (2010), "Identification and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity" *mimeo.*

Fernándes, A. and A. Pakes (2008), "Factor Utilization in Indian Manufacturing: A Look at the World Bank Investment Climate Survey Data" NBER Working Paper 14178.

Foster, L., J. Haltiwanger, and C. Syverson (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98(1), 394–425.

Gao, J. (2012), "Identification, Estimation and Specification in Semi-Linear Models". In Handbook on Non- and Semi-Parametric Econometrics and Statistics (edited by Jeffrey Racine, Liangjun Su and Aman Ullah). Oxford University Press.

Henderson, D., R. Carroll and Q. Li (2008), "Nonparametric Estimation and Testing of Fixed Effects Panel Data Models" *Journal of Econometrics* 144, 257–275

Huang G. and Hu Y. (2011), "Estimating production functions with robustness against errors in the proxy variables" CeMMAP working papers CWP35/1

Klette, T. J. and Griliches, Z. (1996), "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous" J*ournal of Applied Econometrics* 11, 343–361.

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

Li, Q. and T. Stengos (1996), "Semiparametric Estimation of Partially Linear Panel Data Models" *Journal of Econometrics* 71, 389–397.

Li, Q. and Sun Y. (2013), "A consistent nonparametric test ofparametric regression functional form in fixed effects panel data models" *Journal of Econometrics*, forthcoming.

Mairesse and Jaumandreu (2005), "Panel-data Estimates of the Production Function and the Revenue Function" *Scandinavian Journal of Economics* 107(4), 651–672.

Olley, S. and Pakes, A.(1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry" *Econometrica* 64, 1263–97.

Ornaghi C.(2006), "Assessing the Effects of Measurement Errors on the Estimation of a Production Function" *Journal of Applied Econometrics* 21, 879–891.

Pakes, A. (1994), "The estimation of dynamic structural models" J. Laffont and C. Sims eds. Advances in Econonometrics, Ch. 5.

Racine, J., & Li, Q. (2004), "Nonparametric estimation of regression functions with both categorical and continuous data" *Journal of Econometrics* 119, 99–130

Roberts, M. and D. Supina (2000), "Output Price and Markup Dispersion in Producer Micro Data: The Roles of Producer Heterogeneity and Noise" in Advances in Applied

Microeconomics, Vol 9, Industrial Organization, Michael Baye (ed.), JAI Press.

Su, L. and Lu, X. (2013), “Nonparametric dynamic panel data models: kernel estimation and specificaton testing” *Journal of Econometrics* 176, 112–133.

Su, L. and A. Ullah (2006), “Profile likelihood estimation of partially linear panel data models with fixed effects," *Economics Letters* 92(1), 75–81.

Su, L. and A. Ullah (2007), “More efficient estimation of nonparametric panel data models with random effects," *Economics Letters* 92(1), 75–81.

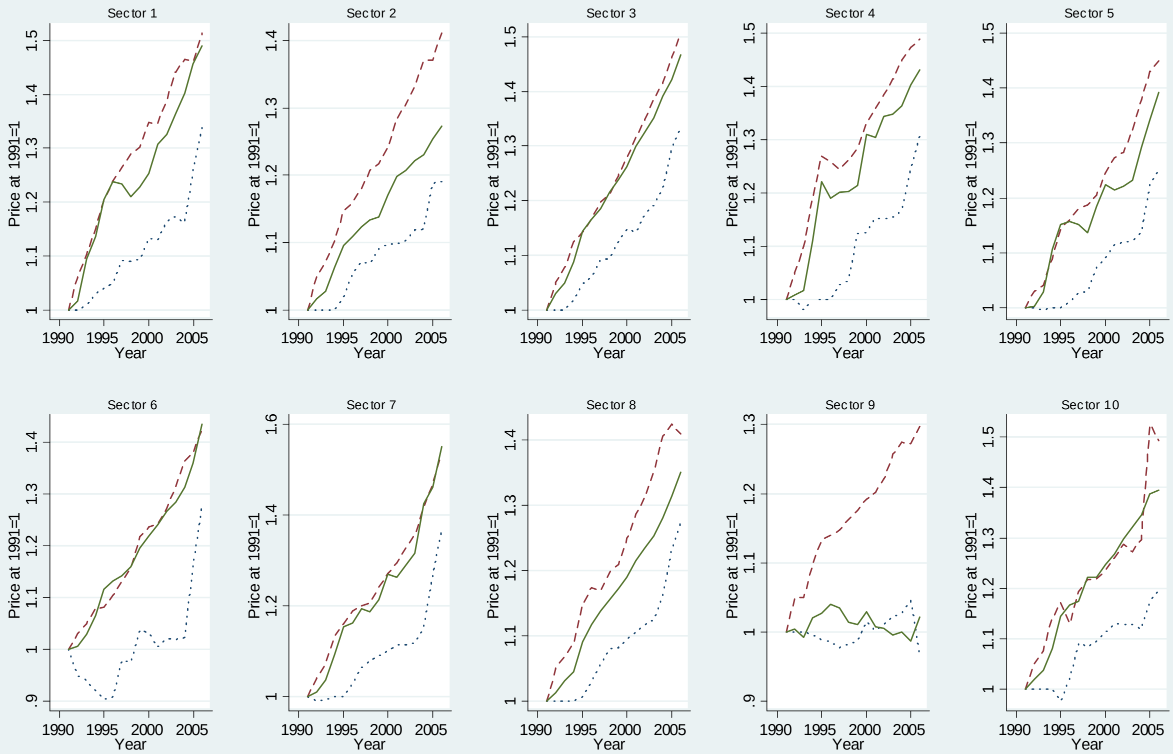
Su, L.and A. Ullah (2010), “Nonparametric and Semiparametric Panel Econometric Models: Specification, Estimation and Testing” *mimeo.*

Syverson, Ch (2011), “What Determines Productivity?” *Journal of Economic Literature* 49(2), 326–65.

Wooldridge, J. (2010), Econometric Analysis of cross section and panel data, MIT press.

Figure 1

Output price index for the 25th and 75th firm and industry price index



Note: dash 75th; solid; Industry PI; dots 25th

Table 1. Unobserved relative price measurement error nonparametric test.

Industrial sectors	I	II	III
1. Food, drink and tobacco	0.157 (0.000)	0.047 (0.000)	0.266 (0.000)
2. Textile, leather and shoes	0.054 (0.000)	0.024 (0.000)	0.081 (0.000)
3. Timber and furniture	0.062 (0.000)	0.036 (0.000)	0.077 (0.000)
4. Paper and printing	0.076 (0.000)	0.038 (0.000)	0.098 (0.000)
5. Chemical products	0.076 (0.000)	0.031 (0.000)	0.109 (0.000)
6. Nonmetalic minerals	0.151 (0.000)	0.072 (0.000)	1.941 (0.000)
7. Metal products	0.050 (0.000)	0.023 (0.000)	0.073 (0.000)
8. Agricultural. & industrial machinery	0.640 (0.000)	0.102 (0.000)	1.400 (0.000)
9. Office, computers and electronics	0.066 (0.000)	0.041 (0.000)	0.084 (0.000)
10. Vehicles accessories	0.105 (0.000)	0.053 (0.000)	0.1585 (0.000)

Note: I: Same bandwidth as Henderson et al. (2008), $h_z=\sigma_z n^{-1/1+q}$, where z index continuos variables and the Li and Racine and Li (2004) mixed continuous-discrete kernel with $\lambda = \sigma_c n^{-1/1+d}$ where c index discrete variables;.Colunms II, III, bandwith multiply by a constant taking values: 0.5, 1.5; p-value in parenthesis obtained after 500 bootstraps.