

Some Non-Parametric Identification Results using Timing and Information Set Assumptions

Daniel Ackerberg

Jinyong Hahn*

University of Michigan

UCLA

December 31, 2015

PLEASE DO NOT CIRCULATE WITHOUT PERMISSION

Abstract

A recent empirical literature on topics such as production function and demand estimation has addressed potential endogeneity problems through use of a combination of timing assumptions on when endogenous variables are chosen by agents, information set assumptions regarding what unobservables are in agents' information sets at various points in time, and Markov assumptions on the unobservable term. This literature has generally relied on parametric assumptions on the primary structural function of interest (e.g. the production function, the demand function). We consider the application of these identifying assumptions in a non-parametric context, and show how the results of Matzkin (2004) and Imbens and Newey (2009) can be applied to show non-parametric identification of the structural function, assuming a scalar unobservable term. We apply the identification argument in a production function context, showing significant non-Hicksian neutral aspects of productivity shocks across three Chilean datasets.

1 Introduction

In panel data contexts, one often desires to make inferences about the effects of an endogenously chosen variable x_{it} on an outcome variable y_{it} . Since assuming orthogonality between

*All errors are our own.

x_{it} and econometric unobservables seems strong, researchers have looked for weaker assumptions on which to base identification and estimation (we loosely interpret orthogonality here to mean either independence, mean independent, or zero correlation, depending on the situation). One general approach is to, instead of assuming that all unobservables are orthogonal to x_{it} , assume that only a portion of the unobservables are orthogonal to x_{it} . The classic linear fixed effects model is perhaps the best known example of this - the unobservable is divided into two components, a time invariant fixed effect component that can be correlated with the x_{it} 's, and a time varying mean zero component that is assumed uncorrelated with x_{it} . The panel data literature, e.g. Chamberlain (1981), Anderson and Hsiao (1982), Arellano and Bond (1991) and Blundell and Bond (1998, 2000), contains a number of generalizations of this assumption. For example, one can estimate models under a sequential exogeneity assumptions whereby the time varying component of the unobservable is allowed to be correlated with future x_{it} 's. Another is Blundell and Bond (2000), who allow the time varying component of the unobservable to contain an AR(1) process of which, e.g. only the innovation in the AR(1) process is assumed uncorrelated with x_{it} (or alternatively, x_{it+1})

A recent literature focused on estimating production functions in a panel context, i.e. Olley and Pakes (1995), Levinsohn and Petrin (2003), Akerberg, Caves, and Frazer (2015), also address endogeneity with this general strategy, but with a different decomposition of the unobservables. Olley and Pakes (1995) assume that the unobservable causing the endogeneity problem, ω_{it} , follows a non-parametric first order Markov Process, i.e. $\omega_{it} = g(\omega_{it-1}) + \xi_{it}$ where $E[\xi_{it}|\omega_{it-1}] = 0$. To identify the production function coefficient on capital k_{it} , they use the assumption that ξ_{it} (but not ω_{it-1}) is mean independent of k_{it} . Loosely speaking, this allows firms' choices of k_{it} to depend on ω_{it-1} , but not ξ_{it} . Akerberg, Berry, Benkard, and Pakes (2007) describe these as *timing and information set* assumptions, i.e. as assumptions regarding 1) the point in *time* at which the agent chooses x_{it} , and 2) the agents' *information sets* at that point in time. Specifically, one interpretation of this assumption is that k_{it} is chosen by firms at time $t - 1$ (i.e. a time-to-build assumption) and that ξ_{it} is not in firms' information sets at time $t - 1$ (while ω_{it-1} is permitted to be in the firms' information sets at $t - 1$).¹

¹OLS, fixed effects, and more general panel data approaches such as Blundell and Bond (1998) can also be interpreted as making timing and information set assumptions, e.g. OLS typically makes the assumption that

While the timing and information set assumptions of Olley and Pakes have been used heavily in the recent production function literature (the three papers referenced above have over 7000 citations), they have also been used in other contexts. For example some recent work on estimation of demand systems, e.g. Berry, Levinsohn and Pakes (1995), Sweeting (2009), Grennan (2013) and Lee (2013), has used these assumptions to address the problem of endogenously chosen product characteristics and/or price (Berry, Levinsohn, and Pakes discussed, but did not actually apply the idea in their empirical work). For example, in some cases product characteristics take time for a firm to design and or change. Hence it might make sense to assume that while current period product characteristics x_{it} are a function of prior periods demand shocks ω_{it-1} , they are not a function of the innovation component ξ_{it} of the current period demand shock. In summary, these timing/information set assumptions of Olley and Pakes can be thought of as a general approach to dealing with endogeneity problems across a variety of literatures.

This literature using these Olley and Pakes timing and information set assumptions has worked under the assumption that the relationship between x_{it} and y_{it} is parametrically specified, and that the unobservable term enters the model linearly. The goal of this paper is show that, at least under certain assumptions, these assumptions also have identifying power in non-parametric situations. Specifically, we show conditions under which these timing and information set assumptions allow us to identify a non-parametric structural relationship between y_{it} and (x_{it}, ω_{it}) . We make particular use of Matzkin's (2004) "unobserved instruments" identification argument, which is also related to control function approaches (Heckman (1978), Blundell and Smith (1989), Blundell and Powell (2003), and Imbens and Newey (2009)). One important limitation of the results is that they rely on dimensionality (scalar) and monotonicity assumptions on unobservables, but this is a limitation of much of the literature on non-parametric identification when one places no parametric restrictions on the structural function.

Our identification results are directly related to at least two other recent papers. Altonji and Matzkin (2005) also study non-parametric identification in panel situations. They consider

none of the unobservables determining y_{it} are in the agent's information set at the time they choose x_{it} , and fixed effects typically makes the assumption that the only such unobservable in the agent's information set is one that is fixed over time.

non-parametric analogues to fixed and random effects estimators. In their setup, the primary endogeneity problem is generated by an unobservable that is fixed over time. This contrasts with our model that follows the spirit of Olley and Pakes, where the problematic unobservable follows a finite M th order Markov process, with timing and information set assumptions like those described above. It is important to note that while these models are different, neither is a generalization of the other.

Hu and Shum (2013) also consider non-parametric identification in panel settings. Like our paper, the problematic unobservable is assumed to be a scalar and follow a finite M th order Markov process. Their setup is more broad than ours in that they allow the outcome variable y_{it} to have a dynamic effects (i.e. y_{it-1} can structurally cause y_{it}). We only consider models without such a dynamic effect. Because the Hu and Shum model is broader than ours, their identification results could be directly applied to our model. However, for the same reason, our identification conditions are weaker. Specifically, we only require the number of observed time periods T to be one greater than the dimension of the Markov process (i.e. $T \geq M + 1$). Hu and Shum require the stronger assumption that $T \geq 3M + 2$. So unlike Hu and Shum, we can estimate a model with a first order Markov process using only two periods of data. The identification results are also quite different in nature - Hu and Shum's rely on deconvolutions, while ours is based on quantiles.

We apply our results to a production function dataset from Chile that has been used extensively in the literature. The majority of work on production functions assumes that the unobservable (i.e. the "productivity shock") enters the production function in a Hicksian neutral way, i.e. linearly in a Cobb Douglas production function. We estimate more flexible functional form production functions where we do not make such an assumption. We find statistically significant evidence of non-Hicksian neutrality of the productivity shock. Interestingly, across the three distinct industries we consider, there are some common patterns of the non-Hicksian neutral productivity effects. Specifically, in all three datasets the productivity shock interacts with labor input in a negative way, i.e. the Hicksian neutral aspect of the shock is negatively correlated with the labor-augmenting aspect of the shock. Our applied results are also related to recent work by Doraszelski and Jaumandreu (2015), who take a different econometric approach to a similar question. Doraszelski and Jaumandreu allows multiple structural

productivity shocks, but relies on a specific parametric specification of the production function. On the other hand, our approach relies on the assumption of a scalar productivity shock, but can allow a completely general non-parametric form of the production function. Given the distinctiveness of the assumptions, we hope the approaches are complementary - empirical conclusions robust to both approaches and set of assumptions would seem to be more convincing than those using only one.

2 Setup

Our goal is to use panel data on observables $\{x_{it}, y_{it}\}$, $i = 1, \dots, N$, $t = 1, \dots, T$ to identify the structural equation

$$y_{it} = f_t(x_{it}, \omega_{it}) \quad (1)$$

where observables x_{it} and unobservable ω_{it} determine a scalar y_{it} . Note that we allow this static structural equation to change in arbitrary ways over time, but the model is not "dynamic" in the sense that y_{it-1} does not directly determine y_{it} . We consider identification of the structural function f_t under the assumption that $N \rightarrow \infty$ and T is fixed.

We consider a situation where the vector of observables x_{it} is endogenously chosen by an economic agent. We start with our key *timing and information set* assumptions,

Condition 1 (*Information Set*) *The agent's information set at t is*

$$I_{it} = \left\{ \{y_{i\tau}\}_{\tau=1}^t, \{x_{i\tau}\}_{\tau=1}^t, \{\omega_{i\tau}\}_{\tau=1}^t, \{\eta_{i\tau}\}_{\tau=1}^t \right\}$$

Condition 2 (*Timing*) *x_{it} is chosen by the agent at time $t - 1$, i.e. according to*

$$x_{it} = h_t(I_{it-1})$$

These assumptions imply that our economic agents are choosing x_{it} without knowledge of the period T structural unobservable ω_{it} , but with knowledge of ω_{it-1} (and y_{it-1} and x_{it-1} , and histories of these variables). Note that the agent's information set at $t - 1$, I_{it-1} , also includes econometric unobservables η_{it-1} . These are other factors that may affect the agent's payoffs and thus the optimal choice of x_{it} . Note that other than these timing and informational

set assumptions, our model is quite general. One nice attribute of our approach is that we will not need to explicitly specify agents' payoffs for our identification results. For example, $x_{it} = h_t(I_{it-1})$ may be the solution to a dynamic programming problem that would require many other auxiliary assumptions to solve. We will not need to specify h_t , and thus can essentially be completely agnostic about these auxiliary assumptions.

A good example of these types of assumption being used in practice is the widely cited and applied Olley-Pakes (1995) approach to estimating production functions. In this context, y_{it} is output (or revenue), x_{it} are inputs chosen by the firm (e.g. capital, labor, R&D) and ω_{it} is an unobservable "productivity" shock. Typically in this literature, some of the inputs in x_{it} are assumed to be chosen *prior* to the firm learning ω_{it} (see Akerberg, Caves, and Frazer (2015) for more discussion of this). Note that in this case, η_{it-1} could represent factors affecting input and output prices (or those prices themselves if they are competitively set). Typically, such factors will impact optimal choices of x_{it} . As noted in the introduction, these assumptions have also been used to generate identification in the empirical demand literature. In that case, η_{it-1} could represent cost shocks that affect firms' choices of product characteristics and prices.

For our non-parametric identification argument we need additional assumptions, specifically on the structural unobservable ω_{it} .

Condition 3 (*Scalar structural unobservable*) $\omega_{it} \in \mathbb{R}^1$

Condition 4 (*Strict monotonicity of structural function*) The inverse function $\omega_{it} = f_t^{-1}(x_{it}, y_{it})$ exists

Condition 5 (*Mth order Markov process*)² $p_t(\omega_{it} \mid I_{it-1}) = p_t(\omega_{it} \mid \{\omega_{i\tau}\}_{\tau=t-M}^{t-1})$ where $M \leq T - 1$

Assumptions (3) and (4) are quite strong. However, scalar and strict monotonicity assumptions are assumptions that are commonly needed in the non-parametric identification literature when one treats the structural function f completely non-parametrically. Note that with

²When $M = 0$, we define $\{\omega_{i\tau}\}_{\tau=t-M+1}^t = \emptyset$. Obviously this is not a particularly interesting case, because in this case, our assumptions imply that ω_{it} is independent of x_{it} , and identification of f_t is trivial using Matzkin (1994).

auxiliary data, one could add additional unobservables to the model that are identified in a preliminary stage. For example, Akerberg, Caves, and Frazer (2015) show how, with additional assumptions and data z_{it} , one can identify ϵ_{it} in the model $\widetilde{y}_{it} = f_t(x_{it}, \omega_{it}) + \epsilon_{it}$ in a preliminary stage, hence reducing the model to the one above, i.e. $y_{it} = \widetilde{y}_{it} - \epsilon_{it} = f_t(x_{it}, \omega_{it})$.

Assumption (5) also may be strong. While the distribution of ω_{it} can vary across time and does not need to be specified parametrically, we do assume that ω_{it} evolves "exogenously" in the sense that conditional on ω_{it} and past values of ω_{it} , the distribution of ω_{it+1} does not depend on values of the other variables in the model dated t and earlier. We also assume that ω_{it} follows a finite M th order Markov process. So unlike Arellano and Bond (1991), Blundell and Bond (1998, 2000), and Altonji and Matzkin (2005), our assumption does not allow there to be a component of ω_{it} that is fixed over time (e.g. a fixed or random effect). On the other hand, we do not require the exchangeability assumption of Altonji and Matzkin (2005). As noted in the introduction, our condition on M is weaker than that in Hu and Shum (2013), who require $T \geq 3M + 2$.

Our necessary assumptions on the other econometric unobservables, i.e. the η_{it} , are considerably weaker. We do not need to limit the dimension of η_{it} , the η_{it} 's can be correlated in any way with I_{it-1} (which includes past values of η), and the η_{it} 's can be contemporaneously correlated with ω_{it} . In addition, the distribution of η_{it} can change over time. However, we do need there to be enough variation in η_{it} to generate sufficient variation in $x_{it+1}(I_{it})$ given ω_{it} . This is because the η_{it} will serve as "unobserved instruments" in the sense of Matzkin (2004).

Given Assumption (5), we can without loss of generality write:

$$\omega_{it} = g_t(\{\omega_{i\tau}\}_{\tau=t-M}^{t-1}, \xi_{it}), \quad (2)$$

where ξ_{it} is a scalar unobservable that is independent of I_{it-1} . Also without loss of generality we can make the following normalizations:

Condition 6 (*Normalizations*) ω_{it} , ξ_{it} and every element of η_{it} have $U(0, 1)$ marginal distributions

Before proceeding with our formal identification arguments, we describe the intuition behind identification in this model. This intuition is actually quite simple. Substituting in lagged (2)

into (1) results in

$$y_{it} = f_t(x_{it}, g_t(\{\omega_{i\tau}\}_{\tau=t-M}^{t-1}, \xi_{it}))$$

Assumption (2) implies that x_{it} is chosen as a function of only I_{it-1} , and ξ_{it} is a scalar unobservable that, given Assumption (5) is independent of I_{it-1} . Therefore, x_{it} is independent of ξ_{it} . Assumptions (3) and (4) guarantee that conditioning on M lags of $\{x_{it}, y_{it}\}$ is equivalent to conditioning on M past values of ω_{it} . Hence, conditional on $\{x_{i\tau}, y_{i\tau}\}_{\tau=t-M}^{t-1}$, variation in x_{it} that is independent of ξ_{it} can be used to identify aspects of f_t .

3 Warm-Up

Consider the linear model with $M = 1$

$$y_t = x_t\beta + \omega_t$$

$M = 1$ implies

$$E[\omega_t | I_{t-1}] = E[\omega_t | y_{t-1}, x_{t-1}, \omega_{t-1}, \eta_{t-1}] = E[\omega_t | \omega_{t-1}] = E[\omega_t | y_{t-1}, x_{t-1}]$$

The last equality follows because of strict monotonicity (we assume the the agent knows the environment, i.e. the structural functions f_t , h_t , and g_t)

Define the random variable

$$\xi_t = \omega_t - E[\omega_t | I_{t-1}] = \omega_t - E[\omega_t | y_{t-1}, x_{t-1}, \omega_{t-1}, \eta_{t-1}] = \omega_t - E[\omega_t | \omega_{t-1}] = \omega_t - E[\omega_t | y_{t-1}, x_{t-1}]$$

By construction,

$$E[\xi_t | I_{t-1}] = 0$$

Next define the random variable

$$\begin{aligned} \varsigma_t &= x_t - E[x_t | y_{t-1}, x_{t-1}] \\ &= h_t(I_{t-1}) - E[h_t(I_{t-1}) | y_{t-1}, x_{t-1}] \end{aligned}$$

The second line follows because our timing assumption implies that $x_t = h_t(I_{t-1})$. Note that variation in ς_t is due to variation in η_{t-1} , i.e. if $\eta_{t-1} = \emptyset$, then $\varsigma_t = 0$ with probability 1. Also note that ς_t is a deterministic function of I_{t-1} .

These results imply that

$$E[\varsigma_t \xi_t \mid I_{t-1}] = \varsigma_t E[\xi_t \mid I_{t-1}] = \varsigma_t * 0 = 0$$

which implies

$$E[\varsigma_t \xi_t \mid y_{t-1}, x_{t-1}] = 0$$

which by definition implies

$$E[(x_t - E[x_t \mid y_{t-1}, x_{t-1}]) (\omega_t - E[\omega_t \mid y_{t-1}, x_{t-1}]) \mid y_{t-1}, x_{t-1}] = 0$$

i.e. that x_t and ω_t are uncorrelated conditional on y_{t-1} and x_{t-1} . This implies that β can be identified by looking at the correlation between x_t and y_t conditional on x_{t-1} and y_{t-1} .

4 Control Function Approach

Returning to the non-parametric case, focus attention on one particular $t \geq M + 1$. Define the random variable

$$\varsigma_t^1 = F_{x_t^1 \mid \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}}^{-1} (x_t^1, \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$$

Now, we consider the second element of x_t conditional on $\{y_\tau\}_{\tau=t-M}^{t-1}$, $\{x_\tau\}_{\tau=t-M}^{t-1}$, and ς_t^1 , i.e., $F_{x_t^2 \mid \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}, \varsigma_t^1}$. Define the random variable

$$\varsigma_t^2 = F_{x_t^2 \mid \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}, \varsigma_t^1}^{-1} (x_t^2, \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}, \varsigma_t^1).$$

By iterating this process, we can create $\varsigma_t = (\varsigma_t^1, \dots, \varsigma_t^J)$.

Theorem 7 x_t is independent of ω_t given $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$.

Proof. By Lemma 13, ξ_t , and ς_t are independent of each other given $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$.

Now note that x_t can be written as a function of $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ and ς_t , say $x_t = \varphi((\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}), \varsigma_t)$. Also, since $\omega_s = f_s^{-1}(x_s, y_s)$, we can see that $\omega_t = g_t(\{\omega_\tau\}_{\tau=t-M}^{t-1}, \xi_t)$ can be written as a function of $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ and ξ_t , say $\omega_t = \phi((\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}), \xi_t)$.

■

This allows us to identify $y_t = f_t(x_t, \omega_t)$ using $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ as a control function following Imbens and Newey (2009) (and is also consistent with the result by Matzkin (2004), who showed equivalence between the control function approach and "unobservable instruments"). In short, consider identification of the inverse function $\omega_t = f_t^{-1}(Y_t, X_t)$, which equals $\Pr[f_t(X_t, \omega_t) \leq Y_t]$ because of the monotonicity in ω_t and the normalization that $\omega_t \sim U(0, 1)$. Denoting $v_{t-1} = \{\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}\}$, we have

$$\begin{aligned} \Pr[f_t(X_t, \omega_t) \leq Y_t] &= \int \Pr[f_t(X_t, \omega_t) \leq Y_t \mid V_{t-1} = v_{t-1}] f_V(v_{t-1}) dv_{t-1} \\ &= \int \Pr[f_t(x_t, \omega_t) \leq Y_t \mid V_{t-1} = v_{t-1}, X_t = x_t] f_V(v_{t-1}) dv_{t-1} \\ &= \int \Pr[y_t \leq Y_t \mid V_{t-1} = v_{t-1}, X_t = x_t] f_V(v_{t-1}) dv_{t-1} \end{aligned}$$

since by Theorem 1, ω_t is independent of x_t conditional on v_{t-1} . Loosely speaking, because of this independence condition, conditional on v_{t-1} , fixing X_t is equivalent to conditioning on realizations of $X_t = x_t$ in the data. Clearly, both the functions inside the integral are directly identifiable from the data. Note that the assumptions required for the Imbens and Newey (2009) identification result do require implicit assumptions on the econometric unobservables η_{it} that determine the vector of endogenous variables x_t . For example, Assumption 2 of Imbens and Newey requires that the support of the conditional distribution of V given $X = x$ is equal the entire support of the marginal distribution of V . It is challenging to elucidate the precise assumptions necessary for identification in our context (at least in a useful form, since it depends on the h_t function which is unspecified and might depend, e.g. on a complicated dynamic programming problem), but generally speaking, there needs to be sufficient variation in the unobservables η_t generating the endogenous x_t . For example, one can construct simple examples where the model is not identified when the dimension of η_t is less than the dimension of x_t . In a production function context, this would correspond to a situation where firms have multiple inputs, but only a lesser number of (unobserved to the econometrician) input prices varying across firms.

Lastly, note that it is straightforward to strengthen the timing and information set assumptions in this model. For example, one could alternatively make the timing assumption

Condition 8 (*Timing*) x_{it} is chosen by the agent at time $t - 2$, i.e. according to

$$x_{it} = h_t(I_{it-2})$$

(or alternatively (and equivalently) make the information set assumption that only $\{\omega_{i\tau}\}_{\tau=1}^{t-1}$ is in the agent's information set at t). In this case, instead of using $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ as the control variables, one would use $(\{y_\tau\}_{\tau=t-M}^{t-2}, \{x_\tau\}_{\tau=t-M}^{t-2})$. While the theoretical identification result is the same in this case, estimation based on this stronger assumption is likely to produce more efficient estimates (all else equal, except decreasing M by 1 to make things comparable), since one will have more variation in x_t conditional on the control variables.

5 Application to Production Function Estimation

We apply these identification results to the estimation of production functions. We use the same yearly (1980-1985) Chilean dataset as do Levinsohn and Petrin (2003), Gandhi, Navarro, and Rivers (2015), and others, and focus on three industries - food products (ISIC code 311), Textiles (code 321), and wood products (code 331)³. Levinsohn and Petrin assume a Cobb-Douglas production function and a Hicksian neutral productivity shock, while Gandhi, Navarro, and Rivers use translog production function, though also with a Hicksian neutral productivity shock. The goal here is to investigate the possibility of the productivity shock entering the production function in a non-Hicksian neutral fashion and to quantify that impact, controlling for the endogeneity of input choices. Note that to put this into our non-parametric framework, we do make a stronger assumption regarding the timing of the choice of labor input l_{it} than does Levinsohn and Petrin. Specifically, we assume that l_{it} is chosen by firms at period $t - 1$ (analogous to the assumption of Levinsohn and Petrin (and us) that k_{it} is determined at period $t - 1$). The hope is that labor market frictions (e.g. unions, other government policy, training) make this assumption reasonable. On the other hand, we are more agnostic about other aspects of the labor choice than are Levinsohn and Petrin - for example, they rule out the possibility of firms facing firm-specific, serially correlated labor price shocks, while we allow such shocks.

³Metals (code 381) was too small a dataset for our non-parametric approach to provide stable estimates.

Direct application of the identification strategy described above based on, e.g. kernel estimation of $\Pr [y_t \leq Y_t \mid V_{t-1} = v_{t-1}, X_t = x_t]$ and $f_V(v_{t-1})$ is challenging due to limited sizes of our datasets. We instead use a sieve maximum likelihood strategy based on polynomial approximations. This also allows us to work up slowly, starting with a simple Cobb-Douglas specification, and then moving to more flexible specifications. Specifically, we start with the following Cobb Douglas model where the productivity shock ω_t is restricted to enter in a Hicksian neutral way, i.e.

$$y_{it} = \beta_0 + \beta_t t + \beta_k k_{it} + \beta_l l_{it} + \omega_{it}$$

Next we consider a Cobb-Douglas model where the productivity shock also can have capital or labor augmenting effects, i.e.

$$y_{it} = \beta_1 + \beta_t t + (\beta_k + \sigma_k \omega_{it}) k_{it} + (\beta_l + \sigma_l \omega_{it}) l_{it} + \omega_{it}$$

In this specification, the parameters σ_k and σ_l measure the non-Hicksian neutral effects of the productivity shock.

Lastly, we add higher order terms in k_{it} and l_{it} , which gives us a Translog model with a non-Hicksian neutral productivity shock.

$$y_{it} = \beta_1 + \beta_t t + (\beta_k + \sigma_k \omega_{it}) k_{it} + (\beta_l + \sigma_l \omega_{it}) l_{it} + (\beta_{kk} + \sigma_{kk} \omega_{it}) k_{it}^2 + (\beta_{kl} + \sigma_{kl} \omega_{it}) k_{it} l_{it} + (\beta_{ll} + \sigma_{ll} \omega_{it}) l_{it}^2 + \omega_{it}$$

Again, the σ parameters measure the extent to which the productivity shock has non-Hicksian neutral effects.

Note that the last two models do not satisfy our strict monotonicity assumption for all values of the parameters. In the second model, for example, strict monotonicity requires that

$$1 + \sigma_k k_{it} + \sigma_l l_{it} > 0 \quad \forall i, t$$

and in the third model, it requires that

$$1 + \sigma_k k_{it} + \sigma_l l_{it} + \sigma_{kk} k_{it}^2 + \sigma_{kl} k_{it} l_{it} + \sigma_{ll} l_{it}^2 > 0 \quad \forall i, t$$

However, in our estimation routines, we did not have problems with our non-linear searches ending up in problematic parts of the parameter space. Hence, we were able to estimated the

models without formally enforcing these restrictions on the parameters, and our final estimates are such that the strict monotonicity assumption holds (and is not binding) for all i and t .

To formally estimate these models, we use conditional maximum likelihood, assuming an AR(1) process for the productivity shock

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it}$$

where ξ_{it} is assumed to be normally distributed. Given our identification arguments, we could allow for more general first order Markov Processes, higher order Markov Processes, and/or non-normal innovations (e.g. mixtures of normals), but given the limited number of observations we wanted to focus our "non-parametric" flexibility on the primary structural component of the model, i.e. the production function. For the first model, we then have

$$\begin{aligned} y_{it} &= \beta_0 + \beta_t t + \beta_k k_{it} + \beta_l l_{it} + \rho\omega_{it-1} + \xi_{it} \\ &= \beta_0 + \beta_t t + \beta_k k_{it} + \beta_l l_{it} + \rho(y_{it} - \beta_0 - \beta_t t - \beta_k k_{it} - \beta_l l_{it}) + \xi_{it} \end{aligned}$$

which we estimate by maximum likelihood based on the premise of our identification assumptions and strategy that the innovation ξ_{it} is independent of all the right hand side variables. The same idea holds for the other two models. In the second model, we have

$$\begin{aligned} y_{it} &= \beta_1 + \beta_t t + (\beta_k + \sigma_k \omega_{it})k_{it} + (\beta_l + \sigma_l \omega_{it})l_{it} + \omega_{it} \\ &= \beta_1 + \beta_t t + (\beta_k + \sigma_k \left(\rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1}}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1})} \right] + \xi_{it} \right))k_{it} \\ &\quad + (\beta_l + \sigma_l \left(\rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1}}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1})} \right] + \xi_{it} \right))l_{it} \\ &\quad + \rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1}}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1})} \right] + \xi_{it} \end{aligned}$$

and in the third model, we have

$$\begin{aligned}
y_{it} &= \beta_1 + \beta_t t + (\beta_k + \sigma_k (\rho \omega_{it-1} + \xi_{it})) k_{it} + (\beta_l + \sigma_l (\rho \omega_{it-1} + \xi_{it})) l_{it} + (\beta_{kk} + \sigma_{kk} (\rho \omega_{it-1} + \xi_{it})) k_{it}^2 \\
&\quad + (\beta_{kl} + \sigma_{kl} (\rho \omega_{it-1} + \xi_{it})) k_{it} l_{it} + (\beta_{ll} + \sigma_{ll} (\rho \omega_{it-1} + \xi_{it})) l_{it}^2 + \rho \omega_{it-1} + \xi_{it} \\
&= \beta_1 + \beta_t t + (\beta_k + \sigma_k \left(\rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1} - \beta_{kk} k_{it-1}^2 - \beta_{kl} k_{it-1} l_{it-1} - \beta_{ll} l_{it-1}^2}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1} + \sigma_{kk} k_{it-1}^2 + \sigma_{kl} k_{it-1} l_{it-1} + \sigma_{ll} l_{it-1}^2)} \right] + \xi_{it} \right)) k_{it} \\
&\quad + (\beta_l + \sigma_l \left(\rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1} - \beta_{kk} k_{it-1}^2 - \beta_{kl} k_{it-1} l_{it-1} - \beta_{ll} l_{it-1}^2}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1} + \sigma_{kk} k_{it-1}^2 + \sigma_{kl} k_{it-1} l_{it-1} + \sigma_{ll} l_{it-1}^2)} \right] + \xi_{it} \right)) l_{it} \\
&\quad + (\beta_{kk} + \sigma_{kk} \left(\rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1} - \beta_{kk} k_{it-1}^2 - \beta_{kl} k_{it-1} l_{it-1} - \beta_{ll} l_{it-1}^2}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1} + \sigma_{kk} k_{it-1}^2 + \sigma_{kl} k_{it-1} l_{it-1} + \sigma_{ll} l_{it-1}^2)} \right] + \xi_{it} \right)) k_{it}^2 \\
&\quad + (\beta_{kl} + \sigma_{kl} \left(\rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1} - \beta_{kk} k_{it-1}^2 - \beta_{kl} k_{it-1} l_{it-1} - \beta_{ll} l_{it-1}^2}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1} + \sigma_{kk} k_{it-1}^2 + \sigma_{kl} k_{it-1} l_{it-1} + \sigma_{ll} l_{it-1}^2)} \right] + \xi_{it} \right)) k_{it} l_{it} \\
&\quad + (\beta_{ll} + \sigma_{ll} \left(\rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1} - \beta_{kk} k_{it-1}^2 - \beta_{kl} k_{it-1} l_{it-1} - \beta_{ll} l_{it-1}^2}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1} + \sigma_{kk} k_{it-1}^2 + \sigma_{kl} k_{it-1} l_{it-1} + \sigma_{ll} l_{it-1}^2)} \right] + \xi_{it} \right)) l_{it}^2 \\
&\quad + \rho \left[\frac{y_{it-1} - \beta_1 - \beta_t(t-1) - \beta_k k_{it-1} - \beta_l l_{it-1} - \beta_{kk} k_{it-1}^2 - \beta_{kl} k_{it-1} l_{it-1} - \beta_{ll} l_{it-1}^2}{(1 + \sigma_k k_{it-1} + \sigma_l l_{it-1} + \sigma_{kk} k_{it-1}^2 + \sigma_{kl} k_{it-1} l_{it-1} + \sigma_{ll} l_{it-1}^2)} \right] + \xi_{it}
\end{aligned}$$

Given the invertibility condition on ω_{it} (and thus ξ_{it}) holds for the parameter vectors searched over, the conditional likelihood function is straightforward to construct for all models.

Tables 1, 2, and 3 present the results from all the models for our three industries, respectively. In the first column of each table, for comparison purposes, we report simple OLS estimation of a Cobb-Douglas production function ignoring the endogeneity problem. A first observation is that, relative to the OLS results, addressing the possible endogeneity of k_{it} and l_{it} through our timing and information set assumptions moves the estimates of returns to scale in the anticipated direction. In all the specifications, the estimate of returns to scale decreases moving from Column 1 to Column 2, consistent with firms with higher productivity shocks ω_{it} using more inputs, causing a positive bias in the OLS results. In Column 3, where we allow the productivity shock to enter in a non-Hicksian neutral way, we do find significant estimates of σ_k and/or σ_l in all three industries, suggesting that there is statistical evidence of these productivity shocks entering the model non-linearly. It is difficult to see the magnitudes of variance imparted on output from these non-Hicksian neutral effects from the coefficient estimates, since they are multiplying k_{it} and l_{it} respectively. The 4th column assesses this

by reporting, at the sample mean of k_{it} and l_{it} , the relative standard deviation imparted by the non-Hicksian effects (relative to the variance of the Hicksian neutral effect). As can be seen from the column, by this measure the effects are smaller than the Hicksian neutral aspect of the shock, though they are statistically significant. What is perhaps most interesting in the results is the fact that in all three industries, the estimate of σ_l is significantly negative. This means that firms for which the productivity shock has a positive Hicksian neutral effect, the labor augmenting effect of the productivity shock is negative. In other words, firms that are unobservably efficient in a Hicksian neutral sense are unobservably inefficient in a labor augmenting sense. It seems interesting that this is the case in all three industries. That said, while it could say something about the structural way that inherent productivity differences affect firms, it also seems possible that results like this could be driven by measurement error in labor, which is not part of the model (e.g. Fox and Smeets (2011)).

The fifth column of the table presents results from our most general model. Given the multiple places that the inputs enter into this model, it is hard to interpret the individual coefficient estimates. Moreover, at least in the two smaller industries (321 and 331), we appear to be pushing the limits of how "non-parametric" we can get given the relatively large standard errors. However, there are still statistically significant non-Hicksian neutral productivity effects, including in the second order terms, i.e. σ_{kk} , σ_{ll} , and σ_{kl} . To summarize the non-Hicksian neutral effects, we can look at the variance (or standard errors) imparted by the productivity shock on the elasticities of output w.r.t. k_{it} and l_{it} . In this model, those elasticities are given by

$$\begin{aligned}\frac{\partial y_t}{\partial k_t} &= (\beta_k + \sigma_k \omega_t) + 2(\beta_{kk} + \sigma_{kk} \omega_t)k_t + (\beta_{kl} + \sigma_{kl} \omega_t)l_t \\ \frac{\partial y_t}{\partial l_t} &= (\beta_l + \sigma_l \omega_t) + 2(\beta_{ll} + \sigma_{ll} \omega_t)l_t + (\beta_{kl} + \sigma_{kl} \omega_t)k_t\end{aligned}$$

and hence the standard deviation of the variation imparted on these elasticities from the productivity shock are proportional to $\tau_k = \sigma_k + 2\sigma_{kk}k_t + \sigma_{kl}l_t$ and $\tau_l = \sigma_l + 2\sigma_{ll}l_t + \sigma_{kl}k_t$ (with the sign indicating their correlation with the Hicksian neutral effect of the shock). For the three datasets (again at the sample means of l_{it} and k_{it}), the values of (τ_k, τ_l) are (0.081, -0.145), (-0.002, -0.050), and (0.002, -0.028) respectively. This seem consistent with the above results, i.e. the negative correlation between the labor augmenting effect of the shock and the

Hicksian neutral effect of the shock. We conclude that 1) there is evidence of statistically significant non-Hicksian neutral productivity shock effects, and 2) these have an interesting pattern, particularly w.r.t. their labor augmenting aspects.

References

- [1] Akerberg, D., Benkard, L., Berry, S., and Pakes, A. (2007) "Econometric Tools for Analyzing Market Outcomes", *Handbook of Econometrics*. Amsterdam: North-Holland.
- [2] Akerberg, D., Caves, K., and Frazer, G. (2015) "Identification Properties of Recent Production Function Estimators", *Econometrica* 83: 2411-2451
- [3] Altonji, J. and Matzkin, R. (2005) "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors", *Econometrica* 73: 1053-1102
- [4] Anderson T.W. and C. Hsiao (1982) "Formulation and Estimation of Dynamic Models using Panel Data" *Journal of Econometrics* 18: 47-82.
- [5] Arellano, M. and S. Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations" *The Review of Economic Studies* 58: 277-297.
- [6] Arellano, M. and Bover, O. (1995) "Another Look at the Instrumental Variable Estimation of Error Components Models" *Journal of Econometrics*, 68: 29-51
- [7] Berry, S., Levinsohn, J., and Pakes, A. (1995) "Automobile Prices in Market Equilibrium" *Econometrica* 63: 841-890
- [8] Blundell, R. and Bond, S. (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models" *Journal of Econometrics* 87: 115-143.
- [9] Blundell, R. and Bond, S. (2000) "GMM estimation with persistent panel data: an application to production functions", *Econometric Reviews* 19: 321-340
- [10] Blundell, R. and Powell, J. (2003) "Endogeneity in Nonparametric and Semiparametric Regression Models". *Advances in Economics and Econometrics, Theory and Applications, Eight World Congress. Volume II*, ed. by M. Dewatripont, L.P. Hansen, and S.J. Turnovsky. Cambridge University Press, Cambridge.
- [11] Blundell, R. and Smith, R. J. (1989) "Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models", *Review of Economic Studies*, 56: 37-58
- [12] Chamberlain, G. (1982) "Multivariate Regression Models for Panel Data", *Journal of Econometrics* 18: 5-46.
- [13] Doraszelski, U. and Jaumandreu, J. (2015) "Measuring the Bias of Technological Change", mimeo, U. of Pennsylvania
- [14] Fox, J. and Smeets, V. (2011) "Does Input Quality Drive Measured Differences in Firm Productivity" *International Economic Review* 52:961-989
- [15] Gandhi, A. Navarro, S., and Rivers, D. (2014) "On the Identification of Production Functions: How Heterogeneous is Productivity," mimeo, U-Wisconsin-Madison.

- [16] Grennan, M. (2013), "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices", *American Economic Review*, 103
- [17] Heckman, JJ (1978), "Dummy Endogenous Variables in a Simultaneous Equation System" *Econometrica*, 46: 931-959
- [18] Hu, Y., and Shum, M. (2013) "Identifying Dynamic Games with Serially Correlated Unobservables" in *Advances in Econometrics (Volume 31): Structural Econometric Models*, Emerald Publishing
- [19] Imbens, G. and Newey, W. (2009) "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity" *Econometrica*, 77:1481-1512
- [20] Lee, R. (2013) "Vertical Integration and Exclusivity in Platform and Two-Sided Markets" *American Economic Review* 103: 2960-3000
- [21] Levinsohn, J. and Petrin, A. (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies* 70: 317-342
- [22] Matzkin, R.(2003) "Nonparametric Estimation of Non-Additive Random Functions," *Econometrica* 71: 1339-1375.
- [23] Matzkin, R. (2004) "Unobservable Instruments", mimeo, Northwestern U.
- [24] McElroy, M.B. (1987) "Additive General Error Models for Production, Cost, and Derived Demand or Share Systems," *Journal of Political Economy*, 95(4): 737-757.
- [25] Olley, S. and Pakes, A.(1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry" *Econometrica* 64:1263-1295
- [26] Sweeting, A. (2013) "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry" *Econometrica* 81:1763-1803

6 Lemmas

Lemma 9 ς_t^1 is independent of $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$.

Proof. By construction,

$$p(\varsigma_t^1 | \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}) \sim U(0, 1)$$

regardless of the realization of $\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}$. ■

Lemma 10 ξ_t, ς_t^1 , and $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ are independent of each other.

Proof. Since $\varsigma_t^1 = F_{x_t^j | \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}}^{-1}(x_t^1, \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$, and since $x_t = h_t(I_{t-1})$ by Condition 2, we can conclude that the ς_t^1 is a function of I_{t-1} . Therefore, both ς_t^1 and $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ are functions of I_{t-1} . Because ξ_t is independent of I_{t-1} by construction, we can conclude that ξ_t is independent of $(\varsigma_t^1, (\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}))$. By Lemma 9, we have ς_t^1 and $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ independent of each other. We therefore conclude that ξ_t, ς_t^1 , and $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ are independent of each other. ■

Lemma 11 $(\varsigma_t^1, \varsigma_t^2)$ is independent of $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$, and ς_t^1 and ς_t^2 are independent of each other.

Proof. By construction,

$$p(\varsigma_t^2 | \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}, \varsigma_t^1) \sim U(0, 1)$$

regardless of the realization of $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}, \varsigma_t^1)$. By Lemma 9, we know that ς_t^1 is independent of $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$. The conclusion follows from these two observations. ■

Lemma 12 $\xi_t, (\varsigma_t^1, \varsigma_t^2)$, and $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ are independent of each other.

Proof. since $\varsigma_t^2 = F_{x_t^j | \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}, \varsigma_t^1}^{-1}(x_t^2, \{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1}, \varsigma_t^1)$, and since $x_t = h_t(I_{t-1})$ by Condition 2, we can conclude that the ς_t^2 is a function of I_{t-1} . Therefore, both $(\varsigma_t^1, \varsigma_t^2)$ and $\{\omega_\tau\}_{\tau=t-M}^{t-1}$ are functions of I_{t-1} . Because ξ_t is independent of I_{t-1} , we can conclude that ξ_t is independent of $((\varsigma_t^1, \varsigma_t^2), \{\omega_\tau\}_{\tau=t-M}^{t-1})$. By Lemma 11, we have $(\varsigma_t^1, \varsigma_t^2)$ and $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ independent of each other, from which the conclusion follows. ■

Lemma 13 ξ_t and ς_t are independent of each other given $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$.

Proof. By iterating Lemmas 9 - 12, we obtain ξ_t, ς_t , and $(\{y_\tau\}_{\tau=t-M}^{t-1}, \{x_\tau\}_{\tau=t-M}^{t-1})$ are independent of each other, from which the conclusion follows. ■

Industry 311 - Food Products						
	OLS CD	Endogenous CD	Endogenous CD + RC	Implied Relative SD at Mean	Endogenous Translog plus RC	Implied Relative SD at Mean
Constant	2.089 (0.059)	1.873 (0.123)	2.072 (0.124)		3.858 (0.449)	
β_k	0.366 (0.010)	0.465 (0.017)	0.438 (0.016)		-0.424 (0.099)	
β_l	1.027 (0.019)	0.818 (0.031)	0.825 (0.030)		1.922 (0.157)	
β_t	0.012 (0.005)	0.038 (0.013)	0.040 (0.013)		0.031 (0.012)	
ρ		0.545 (0.010)	0.554 (0.010)		0.514 (0.011)	
σ_ξ		0.661 (0.006)	0.617 (0.027)	1	0.712 (0.141)	1
σ_k			0.050 (0.008)	0.432	-0.057 (0.038)	0.491
σ_l			-0.108 (0.013)	-0.360	0.033 (0.080)	0.108
β_{kk}					0.050 (0.009)	
β_{ll}					-0.141 (0.029)	
β_{kl}					-0.005 (0.030)	
σ_{kk}					0.015 (0.005)	1.161
σ_{ll}					0.021 (0.015)	0.251
σ_{kl}					-0.038 (0.018)	1.111

Industry 321 - Textiles						
	OLS CD	Endogenous CD	Endogenous CD + RC	Implied Relative SD at Mean	Endogenous Translog plus RC	Implied Relative SD at Mean
Constant	3.559 (0.103)	3.441 (0.262)	3.446 (0.270)		2.047 (1.104)	
β_k	0.279 (0.018)	0.386 (0.034)	0.386 (0.036)		0.627 (0.279)	
β_l	0.899 (0.030)	0.702 (0.053)	0.686 (0.053)		0.839 (0.314)	
β_t	-0.026 (0.009)	-0.043 (0.025)	-0.032 (0.024)		-0.031 (0.024)	
ρ		0.606 (0.021)	0.608 (0.021)		0.597 (0.022)	
σ_ξ		0.629 (0.010)	0.820 (0.050)	1	1.195 (0.276)	1
σ_k			-0.006 (0.010)	-0.051	-0.096 (0.054)	-0.833
σ_l			-0.052 (0.017)	-0.182	0.016 (0.087)	0.056
β_{kk}					-0.018 (0.022)	
β_{ll}					-0.045 (0.053)	
β_{kl}					0.024 (0.060)	
σ_{kk}					0.011 (0.005)	0.852
σ_{ll}					0.025 (0.016)	0.323
σ_{kl}					-0.027 (0.016)	-0.873

Industry 331 - Wood Products						
	OLS CD	Endogenous CD	Endogenous CD + RC	Implied Relative SD at Mean	Endogenous Translog plus RC	Implied Relative SD at Mean
Constant	3.134 (0.130)	2.976 (0.262)	3.010 (0.279)		4.976 (1.275)	
β_k	0.246 (0.020)	0.274 (0.030)	0.282 (0.033)		-0.238 (0.283)	
β_l	1.014 (0.038)	0.937 (0.054)	0.909 (0.058)		1.117 (0.301)	
β_t	-0.019 (0.011)	0.024 (0.027)	0.022 (0.027)		0.033 (0.027)	
ρ		0.513 (0.024)	0.513 (0.024)		0.506 (0.025)	
σ_ξ		0.742 (0.014)	0.739 (0.075)	1	1.132 (0.395)	1
σ_k			0.029 (0.016)	0.259	-0.144 (0.046)	-1.247
σ_l			-0.075 (0.032)	-0.256	0.178 (0.137)	0.611
β_{kk}					0.004 (0.018)	
β_{ll}					-0.168 (0.042)	
β_{kl}					0.115 (0.044)	
σ_{kk}					0.011 (0.005)	0.897
σ_{ll}					-0.010 (0.249)	-0.125
σ_{kl}					-0.015 (0.024)	-0.488