Timing Assumptions and Efficiency: Empirical Evidence in a Production Function Context

Daniel A. Ackerberg*

September 30, 2016

Abstract

Much of the recent empirical work estimating production functions has used methodologies proposed in two distinct lines of literature: 1) the literature started by Olley and Pakes (1996) on "proxy variable" techniques, and 2) what is commonly referred to as the "dynamic panel" literature. We illustrate how timing and firm information set assumptions are key to both methodologies, and how these assumptions can be strengthened or weakened almost continously. We also discuss other assumptions that have utilized in these literatures to increase the precision of estimates. Empirically, we then examine how, in a number of plant level production datasets, strengthening or weakening the timing/information set assumptions affects the precision of estimates. We compare these impacts on precision to those achieved by imposing other potential assumptions. This gives the researcher a better idea of the efficiency tradeoffs between different possible assumptions in the production function context.

1 Introduction

There is a large and active empirical literature that estimates production functions relating outputs to inputs of firms. It has long been recognized that there are significant econometric hurdles here. In particular there is the issue that observed inputs will likely be endogenously chosen as a function of unobservable components of production, generating endogeneity issues. Classic solutions to these endogeneity problems, e.g. fixed effects approaches or instrumental variables based on observing exogenous shifters of inputs (e.g. observed input prices), have not always been successful at addressing these issues.

Given this, researchers have turned to other methodologies to estimate production functions accounting for endogeneity. In particular, two methods have seen extensive use in the empirical literature. First, a literature starting with Olley and Pakes (1996) and Levinsohn and Petrin (2003) directly addressing production function estimation has seen extensive use - these two papers have more than 7000 citations between them. We call these methodologies "proxy variable" approaches. Second, a broader set of methodologies, building from the panel data literature, has been applied to production function estimation. Blundell and Bond (2000), which deals specifically with production function estimation, is perhaps the most well known paper detailing this approach as applied to production functions- it has

*Dept. of Economics, University of Michigan; Thanks to Jim Tybout for access to the data, and to seminar participants at Columbia and Northwestern Universities.

more than 1500 citations. We describe these methodologies as "dynamic panel" approaches (though this is somewhat of a misnomer, see Section 2.3).

While obviously quite popular, as might be expected from methodologies that address endogeneity problems *without* directly observing exogenous variation (as in standard instrumental variables methodology), applying these methods can be challenging. First, these methodologies need to make what are arguably strong assumptions on firm behavior. Second, there can be issues with precision of estimates - often, obtaining precise estimates requires enforcing stronger assumptions than one might have liked. On the other hand, there are many things to like about these methods - for example, they are both semi-parametric in the sense that structural production function parameters can be estimated without fully specifying large parts of firms decision making problems. Not only can this help avoid misspecification, but since these decision making problems are often thought to be dynamic, they have computational advantages.

In this paper, we discuss the similarities and differences of he various assumptions used in these two literatures. We pay particular attention to *timing* and *information set* assumptions that we argue are key components of *both* of these methodological approaches. We argue that one nice aspect of these timing/information set assumptions is that they are not 0-1 assumptions - instead, they are quite flexible in that they can be strengthened or weakened almost continuously (up to the discreteness of one's data). Moreover, with data being observed at higher and higher frequency, the flexibility one has in terms of these assumptions is increasing. With two commonly used production datasets, we then investigate the effects of strengthening or weakening these assumptions on the precision of one's estimates. To benchmark the magnitude of these effects on precision, we compare them to the effects on precision of alternative auxiliary assumptions that have been made in these literatures. In particular, in the dynamic panel literature, an additional stationarity assumption has often been imposed on the model, also increasing precision. We compare these precision gains, e.g. showing that in some cases, the increased precision one obtains by strengthening ones timing/information set assumption by one period is similar to the increased precision one obtains by adding the stationarity assumption. The general point of the paper is to illustrate is that there are different dimensions in which one can add (or subtract) assumptions in these models, and one should consider which might be the most appropriate or credible in the particular data one has. We note that these timing and information set assumptions have also seen recent application in other literatures, e.g. demand estimation (e.g. Sweeting (2009), Grennan (2013) and Lee (2013)) so our results could be interesting in that context as well.

2 Proxy Variable and Dynamic Panel Approaches

We start by outlining the proxy variable and dynamic panel approaches. We focus on timing and information set assumptions that we argue are common to the two approaches - illustrating how these assumptions can be weakened or strengthened almost continuously. We also examine other assumptions or choices amongst assumptions in the methods - highlighting other similarities and differences across the approaches.

2.1 Proxy Variable Approaches

We first decribe what are often called "proxy variable" approaches, as introduced by Olley and Pakes (1996 - henceforth OP). This has been an active are of recent methodological study - including work by Levinsohn and Petrin (2003 - henceforth LP), Ackerberg, Caves, and Frazer (2006, 2015 - henceforth ACF), Van Biesebroeck (2007), Wooldridge (2009), De Loecker (2011), Doraszelski and Jaumandreu (2013), Gandhi, Navarro and Rivers (2016), Kim, Petrin, and Song (2016), Grieco, Li and Zhang (2016) and Collard-Wexler and De Loecker (2016). Consider a simple Cobb-Douglas production function in logs

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \epsilon_{it} \tag{1}$$

where y_{it} is the log of output, k_{it} is the log of capital input, and l_{it} is the log of labor input, all of which are observed by the econometrician. There are two econometric unobservables, ω_{it} and ϵ_{it}. The ϵ_{it} represent shocks to production or productivity that are not observable (or predictable) by firms before making their input decisions at t. These could be, for example, i.i.d. true shocks to output, or serially correlated measurement error in output. In contrast, the ω_{it} represent "productivity" shocks that *are* potentially observed or (partially) predictable by firms when they make input decisions. ω_{it} could represent variables such as firm managerial ability of a firm or some other unobserved (exogenously evolving) input of production. Note that given these assumptions, ω_{it} (and not ϵ_{it}) is the unobservable that is problematic in terms of econometric endogeneity - given ω_{it} are known or partially known when inputs are chosen, those inputs will generally be correlated with ω_{it}.

In the canonical proxy variable model, ω_{it} is assumed to evolve according to an exogenous first order markov process, i.e.

$$p(\omega_{it}|\, \{\omega_{i\tau}\}_{\tau=0}^{t-1}) = p(\omega_{it}|\omega_{it-1})$$

where the firm knows the distribution p. Typically, firms are assumed to observe ω_{it} at t, but have no information on future ω's (other than their conditional distribution p). Hence

$$p(\omega_{it}|\ I_{it-1}) = p(\omega_{it}|\omega_{it-1})$$

where I_{it-1} is the firms information set at t. Assumptions are also made on the points in time in which the inputs k and l are chosen by the firms. For example, it is often assumed that k_{it} is chosen by the firm at time $t-1$, while l_{it} is chosen at time t. This reflects an economic assumption about how far in advance input decisions need to be made, and in this case represent a "time-to-build" assumption on capital of one period. Given this assumption, Ackerberg, Benkard, Berry, and Pakes (2007) describe capital as a "fixed" input, and labor as a "variable" input.

Estimation in the proxy variable literature typically proceeds in two stages. As discussed in ACF, the first stage can be thought of as a way to identify ϵ_{it}, i.e. one of the two econometric unobservables. Doing this typically relies on an additional assumption that some observed decision variable of the firm at t (e.g. capital investment in OP, some intermediate input in LP) is strictly monotonic (or weakly monotonic) in only one unobservable, ω_{it}. This allows one to invert that decision in the unobservable ω_{it} and thus express that unobservable as a function of data and parameters. Intuitively, given the assumption that ϵ_{it}

is not known to (or irrrelevant for) the firm, ϵ_{it} should not enter these decisions. Of course, as detailed in prior work, the assumption that this decision variable only depends on a single unobservable rules out other unobservables that might affect these observed decision variables, e.g. unobserved input prices.[1]

In the current paper we do not focus on this first stage "inversion". Hence we proceed under the assumption that ϵ_{it} has been identified and netted out of the production function (1). In other words, we consider the simpler model *without* ϵ_{it}, i.e.

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} \tag{2}$$

Of course, one could also get to this point by just starting with a model with no ϵ_{it}, in which case the assumptions required for the first stage inversion would not be necessary.

Estimation of (2), i.e. the production function net of ϵ_{it}, relies on moment conditions resulting directly from the information set and timing assumptions described above. Specifically, because of the first order markov assumption and information set assumptions on ω_{it}, we can decompose ω_{it} into its conditional expectation given I_{it-1} and an innovation term

$$\omega_{it} = E[\omega_{it} \mid I_{it-1}] + \xi_{it} = E[\omega_{it} \mid \omega_{it-1}] + \xi_{it} = g(\omega_{it-1}) + \xi_{it}$$

where by construction, the innovation term ξ_{it} satisfies

$$E[\xi_{it} \mid I_{it-1}] = 0$$

Substituting this decomposition into the production function, we obtain

$$\begin{aligned} y_{it} &= \beta_0 + \beta_k k_{it} + \beta_l l_{it} + g(\omega_{it-1}) + \xi_{it} && (3) \\ &= \beta_0 + \beta_k k_{it} + \beta_l l_{it} + g(y_{it-1} - \beta_0 + \beta_k k_{it-1} + \beta_l l_{it-1}) + \xi_{it} && (4) \end{aligned}$$

and since we can write the innovation term as a function of data and parameters (where the parameters include the g function), i.e.

$$\xi_{it} = y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - g(y_{it-1} - \beta_0 + \beta_k k_{it-1} + \beta_l l_{it-1}$$

is straighforward to construct moment conditions that can be used for estimation, i.e.

$$E[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - g(y_{it-1} - \beta_0 + \beta_k k_{it-1} + \beta_l l_{it-1}) \mid I_{it-1}] = 0$$

Note that under the specific timing assumptions made earlier, $k_{it}, l_{it-1}, y_{it-1} \in I_{it-1}$. Further lags of these variables are also in I_{it-1}, though leads of these variables (in particular, l_{it}) are not. Olley and Pakes (1996) and Pakes and Olley (1995) show that the g function can be treated non-parametrically, though if one assumes that ω_{it} follows an AR(1) process, i.e. $g(\omega_{it-1}) = \rho\omega_{it-1}$, a simple set of moments that would tend to exactly identify the parameters is:

[1] This is discussed further in a later section of this paper.

(5)

$$E\left[\begin{array}{c}(y_{it}-\beta_0-\beta_k k_{it}-\beta_l l_{it}-\rho(y_{it-1}-\beta_0+\beta_k k_{it-1}+\beta_l l_{it-1}))\cdot\begin{pmatrix}k_{it}\\ l_{it-1}\\ y_{it-1}-\beta_0+\beta_k k_{it-1}+\beta_l l_{it-1}\end{pmatrix}\\ y_{it}-\beta_0-\beta_k k_{it}-\beta_l l_{it}\end{array}\right]=0$$

where the last moment enforces a normalization $E[\omega_{it}]=0$ to identify the constant term β_0.

While moments similar to this have been used many times in the literature, exact conditions under which this system identifies the parameters have not been formally derived. This is probably because this procedure implicitly treats large parts of the model non-parametrically (which is a nice attibute of the methodology). In particular, the above model does not fully specify the process by which the firm chooses their inputs - it really only makes timing assumptions about when those inputs are chosen, and what is in firms' information sets when those inputs are chosen (plus the assumptions necessary for the first stage). It is hard to even enumerate the large set of models consistent with these assumptions, let alone show which subset of these models would be identified with this moment condition. That said, there is work that examines various aspects of this issue, e.g. Bond and Söderbom (2005), ACF (2015), and GNR (2016). Some simple intuition on this issue that is apparent from examination of (5) is that identification β_k and β_l will be aided by exogenous variation in k_{it} and l_{it} conditional on $\omega_{it-1}=y_{it-1}-\beta_0+\beta_k k_{it-1}+\beta_l l_{it-1}$. In other words, one should think about why two firms with the same lagged productivity shock would have different values of k_{it} and l_{it}, and why this variation is exogenous. The exogeneity condition is important, for example, because while in this model l_{it} will certainly vary conditional on ω_{it-1} (since unlike k_{it-1}, l_{it} can be chosen after the firm realizes ω_{it}) the part of that variation due to ξ_{it}, i.e. the innovation component of ω_{it}, is obviously not orthogonal to ξ_{it}. Some lessons from the aforementioned literature is that aspects of the model such as dynamics (e.g. adjustment costs) in k_{it} and/or l_{it} or serially correlated (not necessarily observed by the econometrician) input prices can aid in generating this sort of variation.

2.2 Strengthening or Weakening Timing/Information Set Assumptions in Proxy Variable Approaches

The focus of this paper is on the timing and information set assumptions that are fundamental to these proxy variable approaches. In thinking about these assumptions, it is first important to emphasize that the orthogonality conditions arising in these approaches require *both* assumptions on both the time at which inputs are chosen *and* assumptions about what is in firms' information sets at those points in time. To illustrate this, consider a model with just capital, i.e.

(6) $$y_{it}=\beta_0+\beta_k k_{it}+\omega_{it}$$

where ω_{it} again follows a simple AR(1) process $\omega_{it}=\rho\omega_{it-1}+\xi_{it}$ (with $E[\xi_{it}|\ I_{it-1}]=0$). Under 1) the *timing* assumption that k_{it} is decided at $t-1$ and 2) the information set assumption that I_{it-1} that only contains ω_{it-1} (and lags) but not ω_{it}, one obtains that while k_{it} is *not* orthogonal to the entire ω_{it}, it

is orthogonal to *part* of ω_{it}, i.e. the innovation component of ξ_{it}. Of course, this orthogonality relies on both assumptions - for example, even if firms decided k_{it} at $t-1$, if the firm observed ω_{it} a period ahead (i.e. observes ω_{it} at $t-1$) , this orthogonality condition would not hold. For that matter, even if firms chose capital levels 10 periods in advance, as long as firms also observed ω_{it}'s 10 periods ahead, the orthogonality condition would again not hold.

Clearly, what is crucial here is the *difference* between the point in time at which k_{it} is chosen and the point in time at which ω_{it} enters the firms information set. In particular, at least with respect to the moments we are considering to identify the parameters, a model in which k_{it} is chosen five periods in advance and ω_{it} is observed four periods in advance is equivalent to a model in which k_{it} is chosen one period in advance and ω_{it} is observed at t. We will denote this difference as Δ, and given a Δ describe our timing/information set assumption as: *At the point in time when the firm chooses k_{it}, the firms information set includes $\{\omega_{i\tau}\}_{\tau=0}^{t+\Delta}$ but does not include $\{\omega_{i\tau}\}_{\tau=t+\Delta+1}^{\infty}$.* In the model above where k_{it} is the only input, is chosen at $t-1$ and ω_{it} is observed at t, $\Delta=-1$. For a given Δ, the model implies the following moment conditions:

$$E[\xi_{it}|\ \{k_{i\tau}\}_{\tau=0}^{t-\Delta-1}]=0 \tag{7}$$

In words, the innovation component of ω_{it} is orthogonal to inputs that were chosen prior to the firm observing ω_{it}.

Given this setup, we can easily think about strengthening or weakening the timing/information set assumption. First, we can strengthen the assumption by decreasing Δ. This corresponds to either assuming that k_{it} is chosen at points in time further in the past, or assuming that ω_{it} is not observed until later points in time. Examination of (7) indicates why decreasing Δ corresponds to a *strengthening* of the timing/information set assumptions - as Δ decreases, the period t innovation is orthogonal to a larger set of k's. Econometrically, this strengthened assumption will generally lead to a more efficient estimator. Perhaps a more intuitive way of illustrating this is through recursive substitution of the AR(1) process into the production function (6), i.e.

$$\begin{aligned}
y_{it} &= \beta_0+\beta_k k_{it}+\omega_{it}\\
&= \beta_0+\beta_k k_{it}+\rho\omega_{it-1}+\xi_{it}\\
&= \beta_0+\beta_k k_{it}+\rho^2\omega_{it-2}+\rho\xi_{it-1}+\xi_{it}\\
&= \beta_0+\beta_k k_{it}+\rho^3\omega_{it-3}+\rho^2\xi_{it-2}+\rho\xi_{it-1}+\xi_{it}\\
&= \beta_0+\beta_k k_{it}+\rho^4\omega_{it-4}+\rho^3\xi_{it-3}+\rho^2\xi_{it-2}+\rho\xi_{it-1}+\xi_{it}\\
&\ \ .\\
&\ \ .
\end{aligned}$$

For a given assumed Δ, consider the $-\Delta$th iterate of the recursive substitution, i.e.

$$y_{it}=\beta_0+\beta_k k_{it}+\rho^{-\Delta}\omega_{it+\Delta}+\sum_{\tau=1}^{-\Delta}\rho^{\tau-1}\xi_{it-\tau}$$

Under the timing/information set assumption Δ, k_{it} is correlated with the first component of the residual $\left(\rho^{-\Delta}\omega_{it+\Delta}\right)$ but orthogonal to the second component of the residual $\left(\sum_{\tau=1}^{-\Delta}\rho^{\tau-1}\xi_{it-\tau}\right)$. As Δ decreases, the first component of the residual becomes smaller in terms of variance, while the second component of the residual becomes bigger. In other words, as Δ decreases, k_{it} is assumed to be orthogonal to "more" of the residual, i.e. a stronger assumption. Note that as $\Delta \to -\infty$ (e.g. if k_{it} is chosen without knowledge of any of the ω's), then we are assuming that k_{it} is orthogonal to the entire residual, and, e.g., OLS would produce consistent estimates of the production function.

In summary, strengthening the timing/information set assumption from $\Delta = -1$ to $-\infty$ corresponds to assuming that k_{it} is orthogonal to a bigger proportion of the residual, and this will generally generate more efficient estimates. We can also weaken the timing/information set assumption by increasing Δ. For example, $\Delta = 3$ coincides with a model where k_{it} is chosen at t and ω_{it} is observed three periods ahead at $t-3$. Interestingly, the implications of increasing Δ above -1 are fundamentally different than decreasing it. This is because as soon as $\Delta = 0$, k_{it} is chosen with full knowledge of ω_{it} and thus is generally correlated with all components of ω_{it}. Hence the proxy variable literature (as well as the dynamic panel literature, see the next section) uses lags of k_{it} that are orthogonal to components of ω_{it}, as shown in (7). For example, when $\Delta = 0$, k_{it-1} is orthogonal to the innovation component of ω_{it}, i.e. ξ_{it}. This distinction between $\Delta \leq -1$ versus $\Delta \geq 0$ is illustrated in Ackerberg and Hahn (2015), who using a very simply method of proof based on verifying conditions in Matzkin (2004) and Newey and Imbens (2009), show conditions under which a non-parametric version of (6), i.e.

$$y_{it} = f\left(k_{it}, \omega_{it}\right)$$

is identified. However, this simple method of proof only applies when $\Delta \leq -1$ (though this doesn't imply that the model is not identified when $\Delta \geq 0$).

Returning to the canonical two input production function of Section 2.1, one can see that in the proxy variable literature, one can choose a different Δ for each input. The moments (5) (as well as the empirical work and monte-carlo experiments in OP, LP, and ACF) reflect the assumption that for capital, $\Delta = -1$, while for labor, $\Delta = 0$. With more inputs, it is straightforward to set up additional moments given whatever assumption on Δ one wants to make for each of the inputs. In the empirical section of this paper, we examine what happens as we vary the strength of the timing assumptions Δ - in particular, we compare to what extent . Thinking precisely about exactly how strong one wants their timing assumptions to be seems particularly relevant given newer high frequency datasets. For example, if one is willing to make the assumption that $\Delta = -1$ in an dataset with annual observations, presumably one should hypothetically be willing to make the assumption that $\Delta = -4$ were the same dataset to instead have quarterly observations, or $\Delta = -12$ in the analagous dataset with monthly observations

2.3 Dynamic Panel Approaches

A second approach that has been used in many recent empirical production function studies is based on the "dynamic panel" literature. This approach stems from a long line of research on panel data models, e.g. Chamberlain (1982), Anderson and Hsiao (1982), Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998). While this is a very general methodology, Blundell and

Bond (2000) do a good job of illustrating how it can be applied to production function estimation. An interesting observation is that much of this literature is based on models with fixed effects and state dependence (i.e. the lagged dependent variable has a causal effect on the current dependent variable). While state dependence is typically *not included* in production function models, the methodology can alternatively be applied to models with no state dependence but a serially correlated unobservable (in addition to a fixed effect).

Blundell and Bond (2000) consider the following production function

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \alpha_i + \omega_{it} + \epsilon_{it} \tag{8}$$

α_i is a firm fixed effect that for now we will allow to be arbitrarily correlated with values of the inputs in all periods. To keep things simple for now, assume ϵ_{it} is measurement error in output that is orthogonal to inputs choices in all periods (we discuss how this assumption can be relaxed later).

The time varying productivity shock ω_{it} is assumed to follow an AR(1) process $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$. Inputs are classified according to their potential correlation with ω_{it} and its components, e.g. ξ_{it}. A period t input is described as *exogenous* if it is orthogonal to ω_{it}. It is described as *predetermined* if it is orthogonal to ξ_{it}, the innovation in ω_{it}. Lastly, it is described as *endogenous* if it is potentially correlated with ξ_{it}, but orthogonal to future period innovations in ω, e.g. ξ_{it}. This classification can be interpreted in terms of the timing and information set assumptions discussed earlier, i.e. assumptions regarding which ξ_{it}'s the firm observes when making their period t input choices. Specifically, the assumption that an input is *exogenous* corresponds to $\Delta = -\infty$, i.e. where the firm observes no ξ_{it}'s when choosing the input; the assumption that an input is *predetermined* corresponds to $\Delta = -1$, i.e. where the firm observes past ξ_{it}'s but not the current ξ_{it} when choosing the input; and the assumption that an input is *endogenous* corresponds to $\Delta = 0$, where the firm observes current ξ_{it} but not future ξ_{it}'s when choosing the input. As we will see momentarily, it is also simple to consider Δ's other than 0, -1, and -∞ in the dynamic panel framework.

Estimating (8) typically involves double differencing - one difference is to isolate the innovation in ω_{it} - the other difference is to eliminate the fixed effect α_i. The intial difference is a ρ first difference, i.e.

$$\begin{aligned} y_{it} - \rho y_{it-1} &= (1-\rho)\beta_0 + \beta_k (k_{it} - \rho k_{it-1}) + \beta_l (l_{it} - \rho l_{it-1}) + (1-\rho)\alpha_i + (\omega_{it} - \rho\omega_{it-1}) + (\epsilon_{it} - \rho\epsilon_{it-1}) \\ &= (1-\rho)\beta_0 + \beta_k (k_{it} - \rho k_{it-1}) + \beta_l (l_{it} - \rho l_{it-1}) + (1-\rho)\alpha_i + \xi_{it} + (\epsilon_{it} - \rho\epsilon_{it-1}) \end{aligned} \tag{9}$$

and the latter difference is a straight first difference, generating

$$(y_{it} - \rho y_{it-1}) - (y_{it-1} - \rho y_{it-2}) = \tag{10}$$

$$\begin{aligned} &= \beta_k [(k_{it} - \rho k_{it-1}) - (k_{it-1} - \rho k_{it-2})] + \beta_l [(l_{it} - \rho l_{it-1}) - (l_{it-1} - \rho l_{it-2})] \\ &\quad + \xi_{it} - \xi_{it-1} + (\epsilon_{it} - \rho\epsilon_{it-1}) - (\epsilon_{it-1} - \rho\epsilon_{it-2}) \end{aligned}$$

The residual in (10) only contains ξ's and ϵ's, and given the above timing and information set assumptions,

will be orthogonal to lags (or leads) of k and l. For example, if $\Delta = -1$ for both k and l, the moment conditions

$$E[\xi_{it} - \xi_{it-1} + (\epsilon_{it} - \rho\epsilon_{it-1}) - (\epsilon_{it-1} - \rho\epsilon_{it-2}) \mid k_{it-1}, l_{it-1}] = 0 \tag{11}$$

hold, since the timing/information set assumption implies k_{it-1} and l_{it-1} are chosen before ξ_{it-1} (and ξ_{it} are in the firms information set). For $\Delta = 0$, this moment condition would not hold - it would only hold for k and l lagged to period $t-2$ and prior.

While the dynamic panel literature seems to focus on the cases where $\Delta = -\infty, -1$,and 0, it is simple to see how this methodology can also be applied with alternative timing/information set assumptions.[2] As one weakens the timing assumptions by increasing Δ from 0, (11) holds for only further lagged k and l. As one strengthens the timing assumptions, (11) will hold for more and more k's and l's into the future.

Following the logic of what was done in the prior section, there is another way to think of strengthening the timing information set assumptions. If, for example, $\Delta = -2$ (for both k and l), for the initial difference (9), we can do a ρ^2 *two period difference* instead of a ρ one period difference, i.e.

$$\begin{aligned} y_{it} - \rho^2 y_{it-2} &= \left(1-\rho^2\right)\beta_0 + \beta_k\left(k_{it} - \rho^2 k_{it-2}\right) + \beta_l\left(l_{it} - \rho^2 l_{it-2}\right) + \left(1-\rho^2\right)\alpha_i + \left(\omega_{it} - \rho^2\omega_{it-2}\right) + \left(\epsilon_{it} - \rho^2\epsilon_{it-2}\right) \\ &= \left(1-\rho^2\right)\beta_0 + \beta_k\left(k_{it} - \rho^2 k_{it-2}\right) + \beta_l\left(l_{it} - \rho^2 l_{it-2}\right) + \left(1-\rho^2\right)\alpha_i + \rho\xi_{it-1} + \xi_{it} + \left(\epsilon_{it} - \rho^2\epsilon_{it-2}\right) \end{aligned} \tag{12}$$

Then doing a straight first difference to eliminate the α_i term, we get

$$\left(y_{it} - \rho^2 y_{it-2}\right) - \left(y_{it-1} - \rho^2 y_{it-3}\right) = \tag{13}$$

$$\begin{aligned} &= \beta_k\left[\left(k_{it} - \rho^2 k_{it-2}\right) - \left(k_{it-1} - \rho^2 k_{it-3}\right)\right] + \beta_l\left[\left(l_{it} - \rho^2 l_{it-2}\right) - \left(l_{it-1} - \rho^2 l_{it-3}\right)\right] \\ &\quad + \rho\xi_{it-1} + \xi_{it} - \left(\rho\xi_{it-2} + \xi_{it-1}\right) + \left(\epsilon_{it} - \rho^2\epsilon_{it-2}\right) - \left(\epsilon_{it-1} - \rho^2\epsilon_{it-3}\right) \end{aligned}$$

$$\begin{aligned} &= \beta_k\left[\left(k_{it} - \rho^2 k_{it-2}\right) - \left(k_{it-1} - \rho^2 k_{it-3}\right)\right] + \beta_l\left[\left(l_{it} - \rho^2 l_{it-2}\right) - \left(l_{it-1} - \rho^2 l_{it-3}\right)\right] \\ &\quad + \xi_{it} + (\rho - 1)\,\xi_{it-1} - \rho\xi_{it-2} + \left(\epsilon_{it} - \rho^2\epsilon_{it-2}\right) - \left(\epsilon_{it-1} - \rho^2\epsilon_{it-3}\right) \end{aligned} \tag{14}$$

Given the timing/information set assumption of $\Delta = -2$, the error term in (14) will be orthogonal to k_{it-1}, l_{it-1} and further lags. This can straightforwardly be extended to stronger timing/information set assumptions - the first step being to do a $\rho^{-\Delta}$, $-\Delta$-period difference, the second step being to do a straight first difference to elimate the α_i term.

[2] The popular Stata command *xtabond2* can accomodate $\Delta = -\infty, -1$,and 0, as well as weaker timing assumptions where $\Delta > 0$. It does not appear to be able to be used for $-\infty < \Delta < -1$ (of course, these can be straightforwardly programmed in Mata or other programming languages). It is also somewhat challenging to enforce the restrictions necessary to consider a model with an AR(1) error term instead of a lagged dependent variable. (see Soderbom (2015)).

2.3.1 Additional Stationarity Assumptions

As noted by Blundell and Bond (1998, 2000), the double differences typically done when the dynamic panel literature is applied to production functions appear to be fairly demanding on the data - specifically, many empirical researchers have found the procedures generate large standard errors, even, for example when $\Delta = -1$ (we do not know of an empirical study that has used this methodology with $\Delta < -1$). Of course, this is not a criticism of the procedures - it is a result of the fact that they are making relatively weak assumptions - i.e. allowing extensive correlation between the unobservable and the explanatory variables. Blundell and Bond (2000) and Bond and Soderbom (2005) do a good job illustrating what aspects of the data generating process tend to help or hinder this imprecision. To mitigate this issue, Blundell and Bond suggest adding some additional assumptions to the model to sharpen predictions - these additional assumptions are based on stationarity restrictions that were also proposed in Arellano and Bover (1995).

In the production function context, these additional assumptions regard how the fixed effects α_i relate to the inputs. In the prior section, the α_i were permitted to be arbitrarily correlated with the inputs. Blundell and Bond (1998, 2000) suggest the additional assumption that the α_i are orthogonal to *specific functions* of the inputs. In particular, they assume that

$$E[\alpha_i|\ k_{it} - k_{it-1}, l_{it} - l_{it-1}] = 0 \tag{15}$$

i.e. that the fixed effects are orthogonal to *changes* in the inputs. In other words, while α_i is permitted to be correlated with levels of inputs, it is assumed to be orthogonal to *changes* in these inputs. Because in these techniques (like the proxy variable techniques), much of the underlying economic model is not specified (e.g. the precise way in which firms make input choices), this is a bit of a high level assumption that can be challenging to interpret. That said, in the production context, this assumption essentially says that inherently more productive firms (i.e. those with high α_i's) are not growing faster (or slower) than those who are less productive. One can think of situations where this assumption might not hold - a primary example would be in a relatively new and growing market, where one might expect better firms to become bigger quicker than worse firms. Relatedly it could be a problem if, in a relatively short panel, industry wide variables are trending a market upwards or downwards. Blundell and Bond (2000) illustrate how in the production function context, (15) can essentially be interpreted as a stationarity assumption on the firm operating environment.

Given this stationarity assumption, one can form additional moments based only on the single ρ-differenced production function, i.e. prior to doing the latter difference to eliminate the α_i. Specifically, under (15), the residual in

$$y_{it} - \rho y_{it-1} = (1-\rho)\,\beta_0 + \beta_k\,(k_{it} - \rho k_{it-1}) + \beta_l\,(l_{it} - \rho l_{it-1}) + (1-\rho)\,\alpha_i + \xi_{it} + (\epsilon_{it} - \rho\epsilon_{it-1}) \tag{16}$$

will be orthogonal to appropriately lagged differences in inputs, e.g. if $\Delta = -1$ then this residual is orthogonal to differences in inputs $k_{it} - k_{it-1}, l_{it} - l_{it-1}$ and prior, while if $\Delta = 0$ then this is orthogonal to differences in inputs $k_{it-1} - k_{it-2}, l_{it-1} - l_{it-2}$ and prior. Note that if one were to strengthen the timing assumption to $\Delta < -1$, then the same logic would apply to form valid moments based on a $\rho^{-\Delta}$,

$-\Delta$-period differenced equation as discussed above. In Blundell and Bond (2000) and following empirical work, these additional moments appear to have be relatively successful at increasing precision of estimates - many empirical papers now use estimators utilizing these additional moments. Estimators that utilize these additional assumptions and moments are often referred to as a "System GMM" (or SYS-GMM) estimator, in contrast to the "Differenced GMM" (or DIFF-GMM) estimator that uses only moments based on, e.g. (10). While the stationarity assumption required to generate these additional moments might be considered strong, it is of course weaker than the implicit assumption that $\alpha_i = 0$ in the proxy variable techniques for estimating production functions, though as we discuss in the next section, there are other dimensions in which the proxy variable literature makes weaker assumptions than the dynamic panel literature.

2.4 Comparison of Assumptions in Proxy Variable and Dynamic Panel Approaches

The discussion above illustrates how timing/information set assumptions are common to both the proxy variable and dynamic panel literatures, and these assumptions can be strengthened or weakened quite easily, essentially in a continuous fashion (up to the discreteness inherent in the data). We now summarize differences in other assumptions across the two methodologies, starting with assumptions on ω_{it}, then proceeding to assumptions on α_i and ϵ_{it} (more discussion on this is in ACF).

Regarding assumptions on ω_{it}, the discussion above illustrates that, generally speaking, the dynamic panel literature makes stronger assumptions than the proxy variable literature. While the proxy variable methodology can straightforwardly allow ω_{it} to follow a general (and potentially non-parametrically estimated) first order markov process $\omega_{it} = g(\omega_{it-1}) + \xi_{it}$, the dynamic panel literature assumes that ω_{it} follows an AR(1) process $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$. The issue is that without that linear structure of the markov process, simple linear differencing as in (9) to isolate the innovation term ξ_{it} is not possible. In contrast, the non-linear structure can be accommodated in the proxy variable literature because of the lack of α_i and the fact that ϵ_{it} is netted out in the first stage. Note that both methodologies can be extended to allow for higher order markov processes - though in the proxy variable approach this requires more data and assumptions regarding the first stage inversion (as discussed in Ackerberg, Berry, Benkard, and Pakes (2007)), and for the dynamic panel approach, simple differencing requires assumption a linear higher order markov process, e.g. $\omega_{it} = \rho_1\omega_{it-1} + \rho_1\omega_{it-2} + \xi_{it}$.

However, while the proxy variable approaches generally makes weaker assumptions on the ω_{it} process, they typically make stronger assumptions regarding α_i. As detailed in the prior section, the dynamic panel literature does allow heterogeneous α_i across firms. In some cases it makes literally no assumptions on the α_i's (when the additional SYS-GMM moments are not used), while in others (when the SYS-GMM moments are used), it makes the stationarity assumptions discussed above.

On the other hand, for the most part the proxy variable literature assumes that $\alpha_i = 0$. It is interesting to note that the primary reason for requiring this assumption is the first stage inversion, which as noted above serves to "identify" ϵ_{it} and net it out of the production function as in (2). This first stage inversion requires that an observed decision variable of the firm (e.g. investment, choice of an input) depend on only a scalar unobservable, i.e. ω_{it}. This scalar unobservable assumption allows one to write ω_{it} as a function of observables and parameters, and this is what allows one to identify ϵ_{it} in the

first stage. The problem with α_i's that vary across firms[3] is that generally speaking, this means that any input decision i will depend on *both* α_i and ω_{it}, hence violating the scalar unobservable assumption and not permitting the first stage inversion. Interestingly, GNR show that in some special cases, one can still do the first stage inversion with α_i's in the proxy variable methodology. Specifically, if the observed decision variable that is being used as the proxy variable depends only on the *sum* $\alpha_i + \omega_{it}$ (and not α_i and ω_{it} individually), then one can still do the first stage inversion. While this condition is not likely to hold for dynamic decision variables like investment (since investment is a long-run decision, it is likely to respond differently to an increase in α_i vs an increase in ω_{it}) , it may often hold for static decision variables, e.g. intermediate inputs.

Lastly, consider assumptions in the two literatures regarding the ϵ_{it}'s. As noted above, simply allowing any ϵ_{it}'s in the proxy variable approach typically requires a host of auxiliary assumptions - in particular the key scalar unobservable and monotonicity assumptions required to do the first stage inversion and net out the ϵ_{it}'s. As noted in ACF, for example, this can require strong assumptions limiting unobserved heterogeneity in demand functions. They describe, for example, how in a value added production function with inputs capital ($\Delta = -1$), labor ($\Delta = 0$), and materials ($\Delta = 0$), that 1) using investment demand (e.g. OP) to proxy for unobserved productivity requires assuming away serially correlated unobserved input price shocks to either capital, labor, or materials, that 2) using unconditional materials demand (e.g. LP) to proxy for unobserved productivity requires assuming away serially correlated unobserved input price shocks to labor and materials, and 3) using conditional materials demand (e.g. ACF - conditional on labor input choice) to proxy for unobserved productivity requires assuming away serially correlated unobserved input prices to materials. The proxy variable approaches also place restrictions on the ϵ_{it} process itself - essentially it requires ϵ_{it} to either be measurement error (potentially serially correlated) in output, or for ϵ_{it} to be real shocks to output but that do not impact choice of the proxy input in any period. For the latter to hold will typically require ϵ_{it} to be independent over time (otherwise optimal input choices will generally depend on past ϵ_{it}) and also rules out past ϵ_{it}'s from impacting future choices of the proxy input, which might, e.g., occur in a context with credit constraints (i.e. current inputs might depend on past shocks to output - see Shenoy (2015) for similar points).

The dynamic panel literature arguably makes weaker assumptions regarding the ϵ_{it}'s. Perhaps most importantly, the dynamic panel literature does not require any first stage inversion to net out the ϵ_{it}'s. Hence it does not need to make the scalar unobservable and monotonicity assumptions that the proxy variable approach does - this means that dynamic panel methods do not place significant restrictions on unobserved, serially correlated, input prices. This can be important, since as observed by, e.g ACF and Bond and Söderbom (2005), serially correlated input prices can be helpful in generating exogenous variaion. The precise assumptions that the dynamic panel literature makes on the ϵ_{it}'s depends on exactly which moments one uses (e.g. DIFF-GMM vs SYS-GMM) and values of Δ. For example, examining (10), the ϵ_{it}'s in the unobserved term will need to be orthogonal to the various contemporaneous or lagged inputs used as instruments. Interestingly, note that, depending on Δ, ϵ_{it}'s could be permitted to be correlated with future (or sufficiently in the future) choices of inputs, which again might occur in a situation with credit constraints. In summary, the dynamic panel literature appears to make stronger

[3] Assuming fixed T - if one assumes $T \to \infty$, allowing α_i's that vary across firms is trivial.

restrictions on the ω_{it} process, while the proxy variable literature tends to make stronger restrictions on the α_i and ϵ_{it}'s.

3 Empirical Work

In the prior sections, we argued that timing and information set assumptions play a key role in both the proxy variable and dynamic panel approaches to production function estimation. Moreover, these timing and information set assumptions can be strengthened or weakened in a very natural, and almost continuous way. The empirical goal of this paper is to examine the relationship between the strengthening or weaking of these assumptions and precision of ones' estimates. Obviously, stronger assumptions will lead to more precise estimates, so for this to be useful, we want to quantify this increased precision in some helpful way. The way we do this is by comparing the increased precision from strengthened timing/information set assumptions to the increased precision from strengthening other assumptions in the production function context. Since, in the dynamic panel literature, many papers utilize the additional SYS-GMM moments (and corresponding stationarity assumptions) proposed by Blundell and Bond (2000), we use this as our benchmark. In other words, we compare the additional precision obtained from strengthening ones timing/information set assumptions vs. the additional precision from moving from the DIFF-GMM moments to the augmented SYS-GMM moments. As noted earlier, one apparent reason for substantial use of the SYS-GMM moments is the lack of precision of estimates generates by DIFF-GMM moments. Our comparisons illustrate how there are alternative ways to increase precision, e.g. strengthing timing assumptions, and our results show, at least in some commonly studied datasets, what the efficiency tradeoff is between the additional SYS-GMM moments and stronger timing assumptions. We feel that, particularly in higher frequency datasets where, e.g. $\Delta < -1$ can still be a short lag, strengthening timing assumptions is something that might be considered.

We investigate this issue in two frequently used production datasets from Chile and Mexico. The Chilean dataset, covering plants with at least 10 workers over 1979-1986, was the subject of LP, for example, and the Mexican plant level data, covering 1984-1990, has also been used by many studies, e.g. Asker, Collard-Wexler, and DeLoecker (2015). For each dataset we focus on 3 industry classifications - food products, clothing, and wood products. These were 3 of the industry classifications considered by LP. In both datasets the period of time is one year, so the distinction between, e.g. $\Delta = -1$ and $\Delta = 0$ is that the former is a stronger timing assumption by a full year.

To compare apples to apples, our empirical results only utilize dynamic panel methodology. Specifically, in the context of the dynamic panel methodology, we investigate and compare the effects of 1) increasing or decreasing Δ and 2) including or not including the additional stationarity moments (i.e. SYS-GMM vs DIFF-GMM). We currently use the STATA command *xtabond2* to compute our estimates. This has a couple of caveats. First, while it is straightforward to allow the inputs to be incrementally more endogenous using the command (i.e. $\Delta = -1$ (predetermined), 0 (endogenous), 1 (more endogenous), 2, 3, ...), it does not appear possible to strengthen the timing assumption, i.e. decreasing Δ below -1. Second, it is not straigtforward to use *xtabond2* to estimate a dynamic panel model with an AR(1) processwith parameter ρ in the error term (i.e. ω_{it} from above). Hence we focus attention on a version

of the model where $\rho = 0$, and consider $\Delta = -1, 0, 1, 2, 3$.[4] Obviously for all the specifications to produce consistent estimates (of parameters and standard errors), one needs that the strongest assumptions hold in all the datasets, i.e. in the data, $\Delta \leq -1$ and the additional SYS-GMM stationarity assumptions to hold.

As our goal is to compare the effect of stronger or weaker assumptions on the precision of production function estimators, our main focus is on the estimated standard errors of our parameter estimates. Being cognizant of the existence of estimation error in the estimates of standard errors (which presumably could be assessed by bootstrapping) and wanting to limit this estimation error, we consider a simple production function with just one input, total labor, and thus one input elasticity coefficient. Table 1 contains our estimates of the *standard errors* of this production function coefficient for the three industries in the two countries. The general patterns of these standard errors across the different estimators are exactly what one would expect. Specifically, as one decreases the strength of the timing assumption from $\Delta = -1$ to $\Delta = 3$, standard errors increase. Similarly, when one moves from the DIFF-GMM specifications to the SYS-GMM specifications that make the additional stationarity assumption, standard errors decrease. To us what is interesting is to compare the magnitude of these standard error changes. In particular, it is interesting to note that starting from a DIFF-GMM specification with $\Delta = 0$, one gets quite similar precision gains from a one period stronger timing assumption (i.e. $\Delta = -1$) as one gets from additionally imposing the SYS-GMM stationarity assumptions. For example, in Chilean Industry 312, the production function coefficient with DIFF-GMM and $\Delta = 0$ has a standard error of 0.157 - moving to $\Delta = -1$ reduces this to 0.049, while adding the stationarity assumptions reduces it to 0.059 (imposing both the additional assumptions together reduces it to 0.023). While the quantitative results will obviously differ across datasets (especially those with different time period lengths), our more general point is that there are different dimensions on which one can strengthen or weaken assumptions in these methodologies, and one should consider all these in trying to find reasonable and credible specifications.

4 Conclusion

We highlight the similarity of two recent literatures on production function estimation in relying on timing and information set assumptions to identify parameters. We show how these assumptions can be weakened or strengthened, and investigate using two production datasets how this changes the precision of the estimates of production function parameters. We compare these precision gains and losses to those from alternative assumptions that have been imposed in the literature, in particular a stationarity assumption often utilized in dynamic panel approaches.

References

Ackerberg, D., Benkard, L., Berry, S., and Pakes, A. (2007) "Econometric Tools for Analyzing Market Outcomes", *Handbook of Econometrics.* Amsterdam: North-Holland.

[4] Given $\rho = 0$, the inability to decrease Δ below -1 is not particularly restrictive, since when $\Delta = -1$, current input choice is already orthogonal to the entirety of ω_{it}.

Ackerberg, D., Caves, K., and Frazer, G. (2015) "Identification Properties of Recent Production Function Estimators" *Econometrica*

Ackerberg, D, and Hahn J (2015) "Some Non-Parametric Identification Results using Timing and Information Set Assumptions" mimeo

Anderson T.W. and C. Hsiao (1982) "Formulation and Estimation of Dynamic Models using Panel Data" *Journal of Econometrics* 18: 47-82.

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.

Arellano, M. and Bover, O. (1995) "Another Look at the Instrumental Variable Estimation of Error Components Models" *Journal of Econometrics*, 68: 29-51

Asker, J., Collard-Wexler, A. and De Loecker, J. (2014) "Dynamic Inputs and Resource (Mis)Allocation" *Journal of Political Economy*

Blundell, R. and Bond, S. (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models" *Journal of Econometrics* 87: 115-143.

Blundell, R. and Bond, S. (2000) "GMM estimation with persistent panel data: an application to production functions", *Econometric Reviews* 19: 321-340

Bond, S. and Söderbom, M. (2005) "Adjustment Costs and the Identification of Cobb-Douglas Production Functions," The Institute for Fiscal Studies, Working Paper Series No. 05/04.

Chamberlain, G. (1982) "Multivariate Regression Models for Panel Data", *Journal of Econometrics* 18: 5-46.

Collard-Wexler, A. and De Loecker, J. (2016) "Production Function Estimation with Measurement Error in Inputs" mimeo

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

De Loecker, J. and Warzynski, F. (2012) "Markups and Firm-level Export Status," *American Economic Review*, 102 (6): 2437-2471.

Doraszelski, U. and Jaumandreu, J. (2013) "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies*, 80: 1338-1383.

Gandhi, A. Navarro, S., and Rivers, D. (2016) "On the Identification of Production Functions: How Heterogeneous is Productivity," mimeo, U-Wisconsin-Madison.

Grennan, M. (2013), "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices", *American Economic Review*, 103

Grieco, P., Li, S., and H. Zhang (2016) "Production Function Estimation with Unobserved Input Price Variation" *International Economic Review*, 57, 2

Griliches, Z. and J.Mairesse. (1998) "Production Functions: The Search for Identification." In: *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, ed. S. Strøm. Cambridge, UK: Cambridge University Press.

Huang, G. and Hu, Y. (2011). "Estimating Production Functions with Robustness Against Errors in the Proxy Variables," mimeo, Johns Hopkins University.

Kim, Kyoo il, Petrin, A. and Song, S. (2016) "Estimating Production Functions with Control Functions when Capital is Measured with Error," *Journal of Econometrics.*

Lee, R. (2013) "Vertical Integration and Exclusivity in Platform and Two-Sided Markets" *American Economic Review* 103: 2960-3000

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

Olley, S. and Pakes, A.(1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry" *Econometrica* 64:1263-1295

Pagan, A. (1986) "Two Stage and Related Estimators and their Applications" *Review of Economic Studies* 517-538

Pakes, A. and Olley, S. (1995). "A Limit Theorem for a Smooth Class of Semiparametric Estimators," *Journal of Econometrics* 65: 295-332.

Shenoy, A. (2016) "Estimating the Production Function when Firms are Constrained", mimeo, UCSC

Söderbom, M (2016) Personal Website http://www.soderbom.net/Resources.htm

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

Van Biesebroeck, J. (2007) "Robustness of productivity estimates," *The Journal of Industrial Economics* , 55: 529-569

Wooldridge (2009) "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters*, 104(3): 112-114.

Table 1

Chile							
			Δ = -1	Δ = 0	Δ = 1	Δ = 2	Δ = 3
Industry 312	Food Products	SYS-GMM	0.023	0.059	0.114	0.123	0.188
n=10790		DIFF-GMM	0.049	0.157	0.356	0.541	0.554
Industry 321	Textiles	SYS-GMM	0.023	0.075	0.109	0.149	0.183
n=2689		DIFF-GMM	0.077	0.144	0.164	0.197	0.216
Industry 331	Wood Products	SYS-GMM	0.036	0.109	0.172	0.281	0.543
n=2539		DIFF-GMM	0.092	0.154	0.292	0.370	0.473
Mexico							
			Δ = -1	Δ = 0	Δ = 1	Δ = 2	Δ = 3
DIV1	Food Products	SYS-GMM	0.042	0.188	0.372	0.290	0.874
n=1849		DIFF-GMM	0.083	0.365	0.775	1.258	2.847
DIV2	Textiles	SYS-GMM	0.052	0.172	0.267	0.360	0.611
n=919		DIFF-GMM	0.078	0.339	0.395	0.839	5.580
DIV3	Wood Products	SYS-GMM	0.106	0.311	0.220	0.489	0.686
n=378		DIFF-GMM	0.224	0.499	1.060	1.270	1.390