IDENTIFICATION PROPERTIES OF RECENT PRODUCTION FUNCTION ESTIMATORS

By Daniel A. Ackerberg, Kevin Caves, and Garth Frazer[1]

This paper examines some of the recent literature on the estimation of production functions. We focus on techniques suggested in two recent papers, Olley and Pakes (1996) and Levinsohn and Petrin (2003). While there are some solid and intuitive identification ideas in these papers, we argue that the techniques can suffer from functional dependence problems. We suggest an alternative approach that is based on the ideas in these papers, but does not suffer from the functional dependence problems and produces consistent estimates under alternative data generating processes for which the original procedures do not.

KEYWORDS: Production functions, productivity, total factor productivity, productivity growth.

1. INTRODUCTION

PRODUCTION FUNCTIONS are a fundamental component of all economics. As such, estimation of production functions has a long history in applied economics, starting in the early 1800s. Unfortunately, this history cannot be deemed an unqualified success, as many of the econometric problems hampering early estimation are still an issue today.

Production functions relate output to inputs (e.g., capital, labor). Perhaps the major econometric issue in estimating production functions is the possibility that there are determinants of production that are unobserved to the econometrician but observed by the firm. If this is the case, and if the observed inputs are chosen as a function of these determinants (as will typically be the case for a profit-maximizing or cost-minimizing firm), then there is an endogeneity problem and OLS estimates of the coefficients on the observed inputs will be biased. The literature has devoted much attention to the problem of these biases and approaches to solving them (see, e.g., Marschak and Andrews (1944), Hoch (1955, 1958, 1962), Mundlak (1961, 1963, 1996), Mundlak and Hoch (1965), Christensen, Jorgenson, and Lau (1973), Christensen and Greene (1976), McElroy (1987), Panzar (1989), and Slade (1989)).

Over the past 20 years, techniques proposed by Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP) to address this endogeneity problem have seen extensive use in the empirical literature, for example, Pavcnik

(2002), Criscuolo and Martin (2009), Fernandes (2007), Blalock and Gertler (2004), Alvarez and Lopez (2005), Sivadasan (2009), Ozler and Yilmaz (2009), Topalova and Khandewal (2011), Kasahara and Rodrique (2008), Dollar, Hallward-Dreimeier, and Mengistae (2005), and numerous others (also see the survey in Ackerberg, Benkard, Berry, and Pakes (2007)). The essence of the OP and LP techniques is that, under certain theoretical and statistical assumptions, one can invert optimal input decisions to essentially allow an econometrician to "observe" unobserved productivity shocks. More precisely, OP identified conditions under which firm-level investment (conditional on capital stock) is a strictly increasing function of a scalar, firm-level, unobserved productivity shock. This strict monotonicity implies that one can invert this investment demand function, and thus "control for" the unobserved productivity shock by conditioning on a nonparametric representation of that inverse function (i.e., a nonparametric function of capital stock and investment). In the first stage of the approach, one estimates the labor coefficient in the production function by regressing output on the labor input and this nonparametric function. LP used a similar approach, but inverted an intermediate input demand function instead of an investment demand function to control for the unobserved productivity shock.

This paper first argues that there may be a problem with the first stage of these estimation procedures. More specifically, under some simple data generating processes that are consistent with the stated assumptions of the OP/LP models, the moment condition underlying the first stage estimating equation *does not* identify the labor coefficient. The problem is one of functional dependence. Under these data generating processes, one can show that labor is a deterministic function of the set of variables that, in the OP/LP procedures, need to be nonparametrically conditioned on. Hence, once one does this nonparametric conditioning, there is no variation in labor left to identify the labor coefficient.[2] We then explore alternative data generating processes, trying to find examples where this functional dependence problem does not apply. We find only a limited set of data generating processes for which the labor coefficient can be identified using the OP/LP first stage moments, characterize these processes, and argue that they may be unrealistic in many industries.

We then propose an alternative estimation procedure that uses moment conditions very similar to those used by OP and LP, but that avoid this functional dependence problem. Specifically, while OP and LP invert investment (OP) and intermediate input (LP) demand functions that are *unconditional* on the labor input, we suggest inverting investment or intermediate demand functions that are *conditional* on the labor input. While use of these conditional input demand functions precludes identification of the labor coefficient using the first stage moments, we describe how this coefficient can be estimated along with

---

[2]In 1999 Susanto Basu made a less formal argument regarding this possible functional dependence problem as a discussant of an earlier version of the Levinsohn–Petrin paper.

the other production function parameters using the second stage moments (after the first stage moments are used to estimate auxiliary parameters). We also illustrate how our use of conditional input demand functions allows for some more general data generating processes than the original OP and LP procedures. Unlike the canonical OP and LP moment conditions, our moment conditions produce consistent estimates if there are unobserved, serially correlated, firm-specific shocks to the price of labor, if labor is chosen prior to other variable inputs (with a different information set), or if labor is dynamic and there are unobserved, firm-specific adjustment costs of labor. Wooldridge (2009) proposed estimating the two sets of moments in LP (or OP) simultaneously, and argued that doing this avoids the same functional dependence issue raised in a working paper version of this paper. On the other hand, as we discuss in Section 4.3.2, the Wooldridge moments as written are based on unconditional input demand functions and do not allow all the aforementioned data generating processes.

We start with a quick review of the various approaches that have been used to address this endogeneity problem over the past 70 years. We then discuss the potential functional dependence issues in OP and LP and describe the data generating processes necessary to avoid these functional dependence issues. We then present our alternative moments and estimation procedure and discuss in what dimensions these alternative moments extend the data generating processes that can be considered. Lastly, we perform a brief Monte Carlo study. We consider some data generating processes that are favorable to the LP moments, and some that are favorable to our alternative moments, and as expected, each procedure performs better in the respective data generating process that is favorable to it. Perhaps more interestingly, we then create misspecification by adding measurement error to the material input variable. Neither our moments nor the LP moments produce consistent estimates, but estimates from our moments are less sensitive than the LP moments to at least this form of misspecification.

2. REVIEW OF LITERATURE ON PRODUCTION FUNCTION ESTIMATION

We start with a brief review of the literature on estimation of production functions (borrowing liberally from Griliches and Mairesse (1998)), ending with a detailed summary of the approaches of Olley and Pakes (1996, henceforth OP) and Levinsohn and Petrin (2003, henceforth LP). OP and LP are the approaches which our work most directly assesses and builds on.

For illustration purposes, consider a simple Cobb–Douglas production function in logs

$$(1) \qquad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it},$$

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. 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. In contrast, the ω_{it} represent "productivity" shocks that *are* potentially observed or predictable by firms when they make input decisions. Intuitively, ω_{it} might represent variables such as the managerial ability of a firm, expected down-time due to machine breakdown, expected defect rates in a manufacturing process, soil quality, or the expected rainfall at a particular farm's location. ε_{it} might represent deviations from expected breakdown, defect, or rainfall amounts in a given year. ε_{it} can also represent (potentially serially correlated) measurement error in the output variable.

The crux of the identification problem inherent in estimating such a production function is that the inputs k_{it} and l_{it} are chosen by firms. This means that if the econometric unobservable ω_{it} is observed (or partially observed) by the firm prior to choosing k_{it} and l_{it}, then these choices will likely depend on ω_{it}, creating correlation between (k_{it}, l_{it}) and ω_{it}, and rendering OLS estimates of β_k and β_l inconsistent. Observation of this endogeneity problem dates back to Marschak and Andrews (1944).

2.1. *Prior Approaches*

One approach to address this endogeneity problem involved using panel data with fixed effects, first advocated by Hoch (1955, 1962), Mundlak (1961, 1963), and Mundlak and Hoch (1965). The most basic of these approaches involves assuming $\omega_{it} = \omega_i$, where the "fixed effect" ω_i is observed by the firm prior to choosing inputs. In essence, the assumption here is that all endogeneity problems are related to a time-invariant shock ω_i. Under this assumption, standard fixed effect approaches can consistently estimate β_k and β_l. Unfortunately, this approach has not worked well in practice; one common finding is unreasonably low estimates of β_k and returns to scale. As noted by Griliches and Mairesse (1998), this could be due to problems with the assumption that the only unobservable generating endogeneity problems is time invariant, or it could be due to problems with data (since differencing can exacerbate measurement error problems; e.g., Griliches and Hausman (1986)). In recent years, these approaches have been generalized (e.g., Chamberlain (1982), Anderson and Hsiao (1982), Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998, 2000))—see Section 4.3.3.

A second approach has utilized information in firms' first order conditions with respect to inputs. Early applications include Klein (1953), Solow (1957), Nerlove (1963), Griliches and Ringstad (1971), and Hall (1988). Generally speaking, these approaches consider estimation of first order conditions regarding, for example, choices of k_{it} and l_{it}. These first order conditions depend on input (and output) prices, as well as the production function parameters (the simplest illustration of this approach is the well-known observation that in a Cobb–Douglas environment, the first order conditions imply that the

cost shares of each input equal the respective input elasticities). These approaches, including the related idea of estimating cost functions (e.g.. Christensen, Jorgenson, and Lau (1973), Christensen and Greene (1976)), have been generalized in a number of dimensions, including some fairly rich theoretical and statistical models. A good example is McElroy (1987), who utilized first order conditions in the context of a production function more general than (1). McElroy not only considered a more flexible production function (translog), but also a more flexible stochastic specification in which there are multiple structural unobservables entering in different places in the production function (instead of only Hicks neutral shocks as in (1)); specifically, there is one such structural unobservable associated with each of the inputs, so, for example, some firms are relatively more efficient at using capital, and some are relatively more efficient at using labor. One caveat of these approaches is that they typically rely on the assumption that input choices satisfy static first order conditions. This could be problematic if certain inputs were subject to significant dynamic issues, for example, adjustment costs, or wedges between purchase and resale prices. A particular worry with respect to this caveat might be the capital input, which in many theoretical models has dynamic implications so that the static first order condition does not hold. In theory, one could construct first order conditions consistent with these dynamic implications, but these would (1) be considerably more complicated than static first order conditions, and (2) likely require additional auxiliary assumptions on the dynamic environment firms operate in.

As pointed out by Griliches and Mairesse (1998), a natural extension of the above "first order condition" approaches involves using input prices as "instruments" for the observed inputs in (1), that is, by simply regressing y_{it} on k_{it} and l_{it}, using p_{it}^k and p_{it}^l as instruments. The idea here is that if input prices are exogenous (i.e., uncorrelated with $\omega_{it} + \varepsilon_{it}$), and if input choices respond to these input prices (which they should), then these input prices provide a source of exogenous variation in k_{it} and l_{it} with which to identify β_k and β_l. What is particularly nice about this simple "input price based IV" method is that its consistency does not require one to specify the relationship between the inputs (k_{it} and l_{it}), and input prices using p_{it}^k and p_{it}^l correctly. Hence, unlike the aforementioned work that relies on using static first order conditions explicitly, it can produce consistent estimates even if the inputs are dynamic in nature (though it will generally not be as efficient if the static first order conditions do hold). Of course, these "input price based IV" methods rely on the researcher observing variation in input prices across firms, and the assumption that these input prices are exogenous. Griliches and Mairesse (1998) noted a number of reasons to potentially question this assumption. For it to hold, for example, the observed price differences across firms cannot capture input quality differences across firms, or different firms' choices of location on a downward sloping input supply curve.

Much of the motivation behind the OP (and LP) approaches is related to the caveats described above. For example, unlike pure fixed effect approaches,

OP and LP do allow for input endogeneity with respect to a time varying unobservable. Unlike many of the first order condition approaches, they can allow for subsets of inputs to be dynamic in nature, yet do not require explicit solution of complicated dynamic first order conditions related to those inputs. And lastly, unlike the "input price based IV" methods, they do not require the econometrician to observe exogenous, across-firm variation in input prices.

Of course, relaxing these assumptions does not come without costs. One needs a new set of statistical and theoretical restrictions to produce identification in the OP and LP models. As detailed extensively below, potentially strong other assumptions need to be made regarding, for example, similarities in the environments that different firms face, what variables are in firms' information sets when different inputs are chosen, and limiting the amount of unobserved heterogeneity in production functions across firms. We see this paper as contributing to the literature by (1) clarifying these statistical and theoretical restrictions, (2) pointing out some significant caveats of these statistical and theoretical restrictions related to functional dependence issues, and (3) proposing some alternative (but very similar) estimators that avoid these functional dependence problems. It is important to point out that while we relax a few of the assumptions of OP/LP and avoid the functional dependence issue, the key assumptions behind our approach remain the same. Our approach, like OP and LP, relies critically on timing assumptions, a scalar unobservable assumption, and a monotonicity assumption. These potentially strong assumptions are what allow for the relaxed assumptions in other dimensions described above. Naturally, whether one chooses our approach or that of the other estimation methods, one should strive for transparency regarding key assumptions and then motivate the appropriateness of those assumptions for the relevant empirical context.

2.2. *Olley/Pakes and Levinsohn/Petrin*

Behind the OP and LP identification strategy is a discrete time model of dynamically optimizing firms. Typically, the primitives of the model are assumed to satisfy the following three assumptions:

ASSUMPTION 1—Information Set: *The firm's information set at t, that is, I_{it}, includes current and past productivity shocks $\{\omega_{i\tau}\}_{\tau=0}^{t}$ but does not include future productivity shocks $\{\omega_{i\tau}\}_{\tau=t+1}^{\infty}$. The transitory shocks ε_{it} satisfy $E[\varepsilon_{it} \mid I_{it}] = 0$.*

ASSUMPTION 2—First Order Markov: *Productivity shocks evolve according to the distribution*

$$p(\omega_{it+1} \mid I_{it}) = p(\omega_{it+1} \mid \omega_{it}).$$

This distribution is known to firms and stochastically increasing in ω_{it}.

ASSUMPTION 3—Timing of Input Choices: *Firms accumulate capital according to*

$$k_{it} = \kappa(k_{it-1}, i_{it-1}),$$

where investment i_{it-1} is chosen in period $t-1$. Labor input l_{it} is non-dynamic and chosen at t.

Assumptions 1 and 2 are straightforward—firms do not observe ω_{it} until time t, but the distribution $p(\omega_{it+1} \mid \omega_{it})$ defines what the firm knows about the distribution of future productivity shocks. Assumption 3 distinguishes the capital and labor inputs. The labor input is non-dynamic in the sense that a firm's choice of labor for period t has no impact on the future profits of the firm. In contrast, capital is assumed to be a dynamic input subject to an investment process. Importantly, Assumption 3 implies that the period-t capital stock of the firm was actually determined at period $t-1$, that is, $k_{it} \in I_{it-1}$. Intuitively, the restriction behind this assumption is that it takes a full period for new capital to be ordered, delivered, and installed. This assumption will be important in generating the moment conditions used for estimation. Neither OP nor LP precisely defined the data generating process (DGP) determining l_{it}. One of the contributions of this paper is to think about this DGP more formally, which we do in Section 3.

In the above model, firms' optimal investment decisions i_{it} are policy functions resulting from a dynamic optimization problem. Under additional assumptions on primitives of the model, OP derived two crucial properties of these policy functions. To save space, we simply state these as additional, high-level, assumptions.

ASSUMPTION 4—Scalar Unobservable: *Firms' investment decisions are given by*

(2) $i_{it} = f_t(k_{it}, \omega_{it}).$

ASSUMPTION 5—Strict Monotonicity: *$f_t(k_{it}, \omega_{it})$ is strictly increasing in ω_{it}.*

Assumption 4 states that investment is a function of the state variables k_{it} and ω_{it} (l_{it} is not a state variable because it is non-dynamic and chosen at t). Griliches and Mairesse (1998) noted that Assumption 4 places strong implicit restrictions on additional firm-specific econometric unobservables in the model. For example, it rules out any unobserved heterogeneity *across firms* in adjustment costs of capital, in demand or labor market conditions (unless they evolve independently over time), or additional unobservables entering other parts of the production function as in McElroy (1987). On the other hand, the

indexing of f by t does allow differences in these variables *across time*.[3] Assumption 5 is a consequence of the assumptions that $p(\omega_{it+1} \mid \omega_{it})$ is stochastically increasing in ω_{it} and that ω_{it} positively affects the marginal product of capital; together, these imply that firms with higher ω_{it}'s have higher expected marginal products of capital in the future, and thus engage in more investment. Formally proving this requires using dynamic programming methods and can be a challenge in more complicated models. Mechanically speaking, all the results here and in later sections can apply to production functions more general than Cobb–Douglas given that Assumptions 1 through 5 hold (so, e.g., they cannot be applied to production functions with multidimensional unobservables like McElroy (1987)).

The econometrician is presumed to observe inputs l_{it} and k_{it}, output y_{it}, and investment levels i_{it} for a panel data set of firms with $N \to \infty$ and fixed T. We assume that inputs and outputs are measured in comparable physical units across observations (extensions to this are discussed in Section 4.3.4). Estimation in OP proceeds in two stages, each stage corresponding to a distinct moment condition. To form the "first stage" moment condition, note that Assumptions 4 and 5 imply that one can invert the investment policy function

$$(3) \qquad \omega_{it} = f_t^{-1}(k_{it}, i_{it})$$

to write the productivity shock as a (unknown) function of observables. Substituting this into the production function, we get

$$(4) \qquad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + f_t^{-1}(k_{it}, i_{it}) + \varepsilon_{it} = \beta_l l_{it} + \Phi_t(k_{it}, i_{it}) + \varepsilon_{it}.$$

Because f_t^{-1} is the solution to a potentially complicated dynamic programming problem, deriving its functional form would require additional computation and specifying additional primitives (e.g., the evolution of industry-wide input prices or demand conditions over time). To avoid this, OP treat f_t^{-1} nonparametrically. This clearly precludes identification of β_0 and β_k using this equation, so these are combined into a composite term $\Phi_t(k_{it}, i_{it})$ that is treated nonparametrically.

The first stage of OP then generates GMM estimates $\widehat{\beta}_l$ and $\widehat{\Phi}_t(k_{it}, i_{it})$ using the moment condition

$$(5) \qquad E[\varepsilon_{it} \mid I_{it}] = E[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, i_{it}) \mid I_{it}] = 0.$$

---

[3]If firm-specific variables such as exogenous labor input prices are observed by the econometrician, they can be included in (2). However, the premise of most of this literature is that such variables are either not available or not believed to be exogenous. If one observed exogenous, across-firm, variation in all input prices, estimating the production function using input price based IV methods might be preferred to OP/LP related methodology (due to fewer auxiliary assumptions). In other words, OP/LP (and the methods we suggest) are probably most useful when one does not observe any exogenous across-firm variation in input prices, or when one observes exogenous across-firm variation in input prices for only a subset of the inputs (for the latter, see, e.g., Doraszelski and Jaumandreu (2013)).

If one uses a polynomial approximation to Φ_t, this can be as simple as running OLS of y_{it} on l_{it} and the polynomial.

For the "second stage" moment condition, Assumptions 1 and 2 imply that we can decompose ω_{it} into its conditional expectation at time $t-1$, and an innovation term, that is,

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

where, by construction, $E[\xi_{it} \mid I_{it-1}] = 0$. Substituting this into the production function gives

$$
\begin{align}
(6) \quad y_{it} &= \beta_0 + \beta_k k_{it} + \beta_l l_{it} + g(\omega_{it-1}) + \xi_{it} + \varepsilon_{it} \\
(7) \quad &= \beta_0 + \beta_k k_{it} + \beta_l l_{it} \\
&\quad + g\big(\Phi_{t-1}(k_{it-1}, i_{it-1}) - \beta_0 - \beta_k k_{it-1}\big) + \xi_{it} + \varepsilon_{it},
\end{align}
$$

where the second line follows from the definition of $\Phi_t(k_{it}, i_{it})$.

Given that $E[\xi_{it} \mid I_{it-1}] = 0$ and $E[\varepsilon_{it} \mid I_{it}] = 0$ (which implies that $E[\varepsilon_{it} \mid I_{it-1}] = 0$), the second stage of OP's estimation procedure uses the following moment condition:

$$
\begin{align}
(8) \quad E[\xi_{it} &+ \varepsilon_{it} \mid I_{it-1}] \\
&= E\big[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} \\
&\quad - g\big(\Phi_{t-1}(k_{it-1}, i_{it-1}) - \beta_0 - \beta_k k_{it-1}\big) \mid I_{it-1}\big] = 0,
\end{align}
$$

where estimation proceeds by "plugging in" the first stage estimates of $\widehat{\beta}_l$ and $\widehat{\Phi}_{t-1}$ into the second stage moment. Note that the two sets of moments, (5) and (8), that identify the production function parameters and the functions Φ_t and g can be thought of in the semiparametric moment restriction framework of Ai and Chen (2003) (and as pointed out by Wooldridge (2009), could be thought of as IV estimators). While these moment conditions could be estimated jointly as in Ai and Chen or Wooldridge (2009), estimation using a two-stage "plug-in" approach is consistent (see Ai and Chen (2007)) and popular in the empirical literature for computational reasons. Joint estimation requires a numeric, nonlinear search over all the parameters of the model (including those representing the functions Φ_t and g), which can be time-consuming and/or unreliable. The two-stage approach requires at most two sequential smaller dimensional searches, and if Φ_t is specified using polynomials or kernels, the first stage estimation of β_l and Φ_{t-1} can be done analytically. On the other hand, as pointed out by Wooldridge (2009), there are advantages to estimating the moment conditions jointly, for example, potential efficiency gains from information in the covariances of the moments (though this is not always the case; see Ackerberg, Chen, Hahn, and Liao (2014)), and more straightforward asymptotic standard errors (see Chen, Hahn, Liao, and Ridder (2013)).

LP take a related approach. The key difference is that instead of using the investment demand equation to "invert" out ω_{it}, they use the demand function for an intermediate input m_{it} (e.g., electricity, fuel, or materials). More specifically, they consider the following production function:

$$(9) \qquad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \varepsilon_{it}$$

and replace Assumptions 4 and 5 with the following:

ASSUMPTION 4B—Scalar Unobservable: *Firms' intermediate input demand is given by*

$$(10) \qquad m_{it} = f_t(k_{it}, \omega_{it}).$$

ASSUMPTION 5B—Strict Monotonicity: $f_t(k_{it}, \omega_{it})$ *is strictly increasing in* ω_{it}.

Assumptions 4b and 5b again allow one to invert ω_{it} as a function of observables, that is, $\omega_{it} = f_t^{-1}(k_{it}, m_{it})$, analogously to the above. Assumption 4b is consistent with a model in which l_{it} and m_{it} are non-dynamic inputs that are chosen simultaneously at t, after the firm has observed ω_{it}; what is crucial is that ω_{it} is the only unobservable entering the intermediate input demand function. Like in OP, this places restrictions on the underlying model, for example, that firms operate in the same labor and material input markets (or identical labor and material input markets), and that they operate in the same output market with either homogeneous goods or completely symmetric product differentiation (or identical output markets). Given these restrictions, Assumption 5b is verified by LP under mild assumptions on primitives.

Other than this, LP proceeds identically to OP. The first stage involves using the equation

$$(11) \qquad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + f_t^{-1}(k_{it}, m_{it}) + \varepsilon_{it}$$

$$(12) \qquad = \beta_l l_{it} + \Phi_t(k_{it}, m_{it}) + \varepsilon_{it}$$

and the moment condition

$$(13) \qquad E[\varepsilon_{it} \mid I_{it}] = E[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, m_{it}) \mid I_{it}] = 0$$

to obtain estimates $\widehat{\beta}_l$ and $\widehat{\Phi}_t(k_{it}, m_{it})$. These estimates are plugged in to the second stage moment condition

$$(14) \qquad E[\xi_{it} + \varepsilon_{it} \mid I_{it-1}]$$
$$= E[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - \beta_m m_{it}$$
$$- g(\Phi_{t-1}(k_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_m m_{it-1}) \mid I_{it-1}] = 0$$

to obtain estimates $\widehat{\beta}_0$, $\widehat{\beta}_k$, and $\widehat{\beta}_m$. LP suggest bootstrapping to obtain standard errors, while Pakes and Olley (1995) provide analytic standard error estimators.

2.3. *Discussion*

As noted by Levinsohn and Petrin, there are a number of advantages to using an intermediate input m_{it} rather than investment i_{it} as a way of "inferring" unobserved productivity. First, theoretically, it can be considerably easier to verify Assumption 5b than Assumption 5; verifying 5 requires analysis of a dynamic programming problem, but as intermediate inputs are typically assumed to be non-dynamic inputs, verifying 5b does not. Second, in practice, investment data can be lumpy; for example, in firm- or plant-level data sets, one often sees $i_{it} = 0$. This casts doubt on Assumption 5 (strict monotonicity), at least for the observations where $i_{it} = 0$. While the OP procedure can be adapted to work in this weakly monotonic situation, it requires discarding data with zero investment, which LP noted can sometimes be a significant portion of the data. Moreover, lumpiness might also occur at points other than $i_{it} = 0$, which is more problematic.

Another important advantage of LP using an intermediate input is related to the critique by Griliches and Mairesse (1998). In OP, Assumption 4 rules out any firm-specific unobservables (other than the productivity shock ω_{it}) affecting investment demand. This, for example, rules out unobserved capital adjustment costs that vary across firms, as well as unobserved, firm-specific shocks to investment prices. The LP method does not rule such shocks out—since the intermediate input and labor are non-dynamic inputs that only affect current profits, there is no reason for m_{it} to depend on these shocks (conditional on k_{it} and ω_{it}). Of course, LP has an analogous restriction regarding unobserved, firm-specific shocks to the price of intermediate inputs and labor. However, this distinction is not symmetric, since OP *also* rules out these material or labor price shocks (except for the special case where they are independent across time). The reason is that if, for example, unobserved wage shocks are serially correlated, the current wage is related to the marginal revenue product of capital in the future, and thus current investment should depend on the current wage, violating Assumption 5. So, in summary, neither OP nor LP allow serially correlated, unobserved heterogeneity (across firms) in prices of labor or intermediate inputs, while only OP rules out unobserved heterogeneity (across firms) in the price of investment or capital adjustment costs.

Lastly, note that one can extend these methods to situations where the econometrician actually *observes* firm-specific variation in the prices of inputs. If these are assumed exogenous, OP/LP can be applied; the prices can simply be included in the appropriate input demand functions above and used as instruments. For example, if exogenous variation in p_{it}^m and p_{it}^l across firms is observed, these variables should be included in the intermediate input equation,

that is, $m_{it} = f_t(k_{it}, \omega_{it}, p_{it}^m, p_{it}^l)$, and included in the instrument set I_{it}. That said, both OP and LP make many auxiliary assumptions, so if one observed exogenous, firm-specific variation in *all* input prices, it might be preferable to use simple "input price based IV" methods, that is, regress y_{it} on k_{it}, l_{it}, and m_{it} using p_{it}^k, p_{it}^l, and p_{it}^m as instruments. This would provide consistent estimates without the OP/LP auxiliary assumptions. Of course, the observed price differences across firms need to be exogenous, that is, the differences cannot capture input quality differences or different firms' choices of location on a downward sloping input supply curve. Having credibly exogenous, observed, variation in the costs of all inputs is fairly rare. Perhaps more likely is a case where one observes exogenous variation in prices of a subset of the instruments. In this case, pure "input price based IV" methods are generally not possible, and we feel that techniques related to OP/LP are valuable. Doraszelski and Jaumandreu (2013) estimated a model where this is the case.[4]

3. FUNCTIONAL DEPENDENCE PROBLEMS

This paper argues that even if the above assumptions hold, there are identification problems with these methodologies. The issue is one of functional dependence in the first step of the respective procedures. We illustrate these issues in the LP context, though one can make very similar arguments in the OP context. To give a brief intuition behind the result, consider a parametric version of the first stage of LP where instead of treating (10) (and its inverse) nonparametrically, one actually uses the parametric FOC corresponding to a Cobb–Douglas production function. The first order condition for m_{it} (conditional on k_{it}, l_{it}, and ω_{it}) is

$$\beta_m K_{it}^{\beta_k} L_{it}^{\beta_l} M_{it}^{\beta_m - 1} e^{\omega_{it}} = \frac{p_m}{p_y},$$

assuming firms are price takers in both input and output markets.[5] Inverting this equation for ω_{it} and substituting into the production function results in

$$(15) \qquad y_{it} = \ln\left(\frac{1}{\beta_m}\right) + \ln\left(\frac{p_m}{p_y}\right) + m_{it} + \varepsilon_{it}.$$

---

[4]In contrast to OP and LP, they treat the inverted input demand function parametrically. They have two good reasons for this. First, including observed input prices in the inverted function increases the dimensionality of the function and makes nonparametric treatment more challenging. Second, since the production function has already been modeled parametrically, one can obtain parametric forms of input demand functions for non-dynamic inputs without significant auxiliary assumptions. Doraszelski and Jaumandreu also extend the OP/LP model to one in which observed firm R&D expenditures affect the evolution of ω_{it}, an important idea that could also be applied in the context of our estimation procedure.

[5]This version of the first order condition assumes that ε_{it} is pure measurement error in y_{it} (if ε_{it} was an unanticipated shock to output, the expectation of $\exp(\varepsilon_{it})$ would enter the first order condition as well).

Since β_l does not enter this expression, a moment condition in ε_{it} based on this equation (which is analogous to the first stage of LP) cannot be informative on β_l.[6]

While the above result might not hold for more general functional forms of the production function, the LP and OP procedures treat the inverted FOCs nonparametrically, which makes identification tougher. The first step moment condition (13) used to estimate β_l in LP is equivalent to the partially linear model studied by Robinson (1988). His condition for identification of β_l (Equation 3.5; see also Ai and Chen (2003), Assumption 3.3 and discussion) is that

$$(16) \qquad E\big[\{l_{it} - E[l_{it} \mid k_{it}, m_{it}, t]\}\{l_{it} - E[l_{it} \mid k_{it}, m_{it}, t]\}'\big] \text{ is positive definite.}$$

An immediate observation is that if l_{it} is functionally dependent on k_{it}, m_{it}, and t (i.e., l_{it} is *only* a function of k_{it}, m_{it}, and t), this condition does not hold. Intuitively, the problem here is that l_{it} is fully determined by the values of k_{it}, m_{it}, and t, so there is no source of variation in l_{it} (conditional on k_{it}, m_{it}, and t) to identify β_l. So we focus our discussion on investigating possible data generating processes for l_{it} and ask the question whether these DGPs generate this functional dependence.

Perhaps the simplest possible DGP for l_{it} is analogous to that for m_{it} (equation (10)), that is,

$$(17) \qquad l_{it} = h_t(k_{it}, \omega_{it}).$$

This assumes that, like m_{it}, labor has no dynamic implications and is chosen with full knowledge of ω_{it}. With this DGP, substituting the inverted (10) into (17) results in

$$l_{it} = g_t\big(k_{it}, f_t^{-1}(k_{it}, m_{it})\big),$$

which states that l_{it} *is* functionally dependent on k_{it}, m_{it}, and t, and implies that the identification condition (16) fails. Intuitively, in equation (12), the contribution of labor to output cannot be separately identified from the nonparametric function $\Phi_t(k_{it}, m_{it})$.

That said, while (17) might be the simplest specification for the data generation process for l_{it}, it is not the only possibility. Our goal now is to search for an alternative DGP for l_{it} (and possibly for m_{it}) that does not suffer this functional dependence problem and is consistent with the basic assumptions of the LP (and OP) procedure(s).

---

[6]The above analysis uses the choice of m_{it} conditional on levels of k_{it}, l_{it}, and ω_{it}. This is most naturally interpreted in the case where m_{it} is chosen after l_{it}. One obtains the same result if one solves simultaneously for optimal choices of both m_{it} and l_{it} conditional on levels of k_{it} and ω_{it}, and plugs the inverted FOC for m_{it} into the production function.

First, consider adding firm-specific input prices to the above model of input choice, for example, prices of labor (p_{it}^l) and materials (p_{it}^m). Generally speaking, both of these firm-specific input prices will affect a firm's choices of both l_{it} and m_{it} (i.e., conditional on p_{it}^m, decreasing p_{it}^l will generally increase a firm's optimal level of *both* l_{it} and m_{it}). Therefore, these input prices would need to be observed by the econometrician, because otherwise (10) would violate the scalar unobservable assumption necessary for the LP inversion. Since intermediate input demand depends on these firm-specific input prices, they end up in the first step nonparametric function, that is,

$$(18) \qquad E[\varepsilon_{it} \mid I_{it}] = E[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, m_{it}, p_{it}^l, p_{it}^m) \mid I_{it}] = 0,$$

and there is again a functional dependence problem, as there are no variables that shift l_{it} but do not enter Φ_t. Another possibility, allowing labor to have dynamic effects, does not help for the same reason. l_{it-1} becomes a state variable of the problem, but since both m_{it} and l_{it} both depend on l_{it-1}, there is still functional dependence.

One way to eliminate the functional dependence problem is to assume that there is "optimization error" in l_{it}. By optimization error, we mean that there is an optimal level of l_{it} (e.g., given by (17)), but that for some reason the firm chooses that optimal level plus noise (that is independent over time and of other errors in the model). One example of this could be workers calling in sick.[7] Such optimization error induces variation in l_{it} conditional on k_{it}, m_{it}, and t, and hence the first step can consistently estimate β_l. However, note that the methodology is not compatible with similar optimization error in m_{it}, since this would violate the scalar unobservable assumption required to invert (10) and obtain ω_{it}. The context of sick days is a good way to illustrate this point with some applied relevance. Suppose that m_{it} are material inputs used in production and that when workers are out sick, less of these materials are used. If the econometrician's measure of m_{it} is the actual level of material input used (which is negatively affected by sick days), there is implicitly optimization error in *both* l_{it} and m_{it}, and because of this optimization error in m_{it}, the first step *will not* produce consistent estimates of β_l. However, if the econometrician's measure of m_{it} is the planned level of material input used (i.e., planned/ordered prior to the realization of sick days, and thus unaffected by sick days), there is only optimization error in l_{it} and the first step *will* produce consistent estimates of β_l. In contrast, the i.i.d. "optimization error" in l_{it} DGP might not be a good representation of a situation where firms' labor choices are distorted by exogenous union issues. This is because union issues are unlikely

---

[7]Note the difference between our "optimization error" and the "optimization error" described by Mundlak (1996). We are simply adding exogenous noise to the desired input level (workers exogenously calling in sick). Mundlak described "optimization error" due to incomplete information; for example, firms choose labor input based on an expected price of labor, but the realized price of labor is different, making their choice of labor input seem suboptimal ex post.

to be i.i.d. over time, and hence either planned or used material inputs are likely to respond to the (unobserved) union shocks, violating the scalar unobservable assumption. Note that classical measurement error in l_{it} (i.e., noise in observed labor that *does not* affect output) is problematic because of standard measurement error problems, and classical measurement error in m_{it} is problematic because it again violates the scalar unobservable assumption that is crucial for the inversion.

A second, similar, DGP that can eliminate the functional dependence problem involves changing the assumption that l_{it} and m_{it} are chosen with the same information set. Consider a point in time, $t-b$, sometime between period $t-1$ and t (i.e., $0 < b < 1$).[8] First, note that if one assumes l_{it} is chosen at $t-b$ and m_{it} is chosen at t, there is a problem because in such a model, optimal m_{it} will generally directly depend on the previously chosen level of l_{it}. Since l_{it} enters (10), the first step moment becomes

$$(19) \qquad E[\varepsilon_{it} \mid I_{it}] = E\big[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, l_{it}, m_{it}) \mid I_{it}\big] = 0,$$

which obviously precludes identification of β_l in the first step. However, one can alternatively consider the reverse situation, that is, where m_{it} is chosen at $t-b$ and l_{it} is chosen at t. Suppose there is some unanticipated firm-specific shock to the price of labor (or demand), $\varkappa_{it}$, that is realized *between* the points $t-b$ and t (i.e., it is not in the firm's information set (I_{it-b}) when m_{it} is chosen, but it *is* in the firm's information set (I_{it}) when l_{it} is chosen). Such a shock will vary l_{it} independently of k_{it}, m_{it}, and t, and thus will eliminate the functional dependence problem. This shock $\varkappa_{it}$ does not need to be observed by the econometrician, but if it is unobserved, it must be i.i.d. over time (otherwise the optimal m_{it+1} will depend on $\varkappa_{it}$, violating the scalar unobservable assumption). Another assumption that is required in this DGP is that the firm observes ω_{it} when choosing m_{it} at time $t-b$ (perhaps more intuitively, that ω does not evolve between time $t-b$ and time t). Otherwise, the nonparametric function of m_{it} and k_{it} will not perfectly control for ω_{it} in the moment condition.

These functional dependence problems (and possible assumptions that avoid them) generalize to the first step of the OP procedure that alternatively uses the investment function to control for productivity variation. Firm-specific input prices (observed or unobserved) and dynamics in l_{it} generally do not break the functional dependence, but optimization error in l_{it} or i.i.d., firm-specific wage (or output price) shocks after i_{it} is chosen do. There is one additional DGP that breaks the functional dependence problem in the OP setup.

---

[8]Note that we continue to assume that production occurs "on the period," that is, at periods $t-1$ and t. Note that this intermediate period, $t-b$, does not need to have a literal interpretation in terms of time. What we are simply trying to do is allow different inputs to be chosen with different information sets, and with this heuristic device, an "input chosen at t" is chosen with more information than is an "input chosen at $t-b$," which is chosen with more information than is an "input chosen at $t-1$."

Continuing with the above setup with the "subperiod" $t - b$, assume that ω evolves between $t - 1$, $t - b$, and t according to a first order Markov process, that is,

$$(20) \qquad p(\omega_{it-b} \mid I_{it-1}) = p(\omega_{it-b} \mid \omega_{it-1})$$

and

$$(21) \qquad p(\omega_{it} \mid I_{it-b}) = p(\omega_{it} \mid \omega_{it-b}).$$

Suppose also that l_{it} is chosen at time $t - b$ with information set I_{it-b}, while i_{it} is chosen at time t with information set I_{it}.

In this case, a firm's optimal investment will follow

$$(22) \qquad i_{it} = f_t(k_{it}, \omega_{it}),$$

while a firm's optimal labor input will not be a function of ω_{it}, but of ω_{it-b}, that is,

$$l_{it} = g_t(\omega_{it-b}, k_{it}).$$

Since ω_{it-b} cannot generally be written as a function of k_{it} and i_{it}, l_{it} will *not* generally be functionally dependent on k_{it} and i_{it}, possibly allowing consistent estimation of β_l in the first step. Note the intuition behind this—labor is chosen without perfect information about what ω_{it} is, and this incomplete information is what generates variation in l_{it} conditional on the nonparametric function $f_t^{-1}(k_{it}, i_{it})$. However, note that this DGP does need to rule out a firm's choice of l_{it} having dynamic implications. If labor did have dynamic effects, then l_{it} would directly impact a firm's choice of i_{it}. As a result, l_{it} would directly enter the first stage nonparametric function and prevent identification of β_l in that first stage.[9]

3.1. *Discussion*

To summarize, it appears that the first stage of the OP and LP procedures correctly identifies β_l only under three very specific DGPs: (1) i.i.d. optimization error in l_{it} (and not in m_{it} (or i_{it})); (2) i.i.d. shocks to the price of labor or output after m_{it} (or i_{it}) is chosen but prior to l_{it} being chosen; or (3) (in the OP context) labor is non-dynamic and chosen at $t - b$ as a function of ω_{it-b}, while

---

[9]Note why this DGP *does not* solve the functional dependence problem in the context of the LP model. In the LP model, if l_{it} is chosen before m_{it}, then m_{it} will *directly* depend on l_{it}, making β_l unidentified in the first stage. In OP, even if l_{it} is chosen before i_{it}, i_{it} does not depend on l_{it} (as long as one maintains the assumption that labor is non-dynamic). This is because i_{it}, unlike m_{it}, is not directly linked to period-t outcomes, and thus l_{it} will not affect a firm's optimal choice of i_{it}.

i_{it} is chosen at t. Note that in practice, one probably would not literally observe the functional dependence problem described above (if one were using polynomials to approximate the nonparametric function, functional dependence would arise in the form of collinearity between l_{it} and the terms in the polynomial approximation). In other words, it is likely that estimation of (5) or (13) will produce an actual numerical estimate of β_l. Our point is that unless one believes that one of these three DGPs is generating the data (and additionally that these DGPs are the *only* reasons why there is no functional dependence), this is simply not a consistent estimator of β_l.

4. OUR ALTERNATIVE PROCEDURE

We now propose an alternative estimation procedure that avoids the functional dependence problems discussed above. Perhaps just as importantly, it also relaxes other assumptions typically made in applications of OP and LP. Specifically, our model allows there to be exogenous, serially correlated, unobserved firm-specific shocks to the price of labor, or firm-specific unobserved adjustment costs to the labor input. It also allows the labor input to have dynamic effects (e.g., hiring or firing costs) more generally. We emphasize that the approach is only a slight adaptation of the OP/LP methodologies, and essentially relies on the same moment conditions. The main difference between our approach and OP and LP is that in our approach, we invert "conditional" rather than "unconditional" input demand functions to control for unobserved productivity. This results in a first stage that *does not* identify the coefficients on variable inputs (e.g., labor). Instead, all coefficients are estimated in the second stage. However, as we shall see, the first stage will still be important to "net out" the untransmitted error ε_{it} from the production function. We start by showing how our method works using the LP intermediate input function before showing the extension to using the OP investment function.

4.1. *The Basic Procedure*

Consider the "value-added" production function

$$(23) \qquad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it}$$

with the following alternative assumptions:

ASSUMPTION 3C—Timing of Input Choices: *Firms accumulate capital according to*

$$k_{it} = \kappa(k_{it-1}, i_{it-1}),$$

where investment i_{it-1} is chosen in period $t - 1$. Labor input l_{it} has potential dynamic implications and is chosen at period t, period $t - 1$, or period $t - b$ (with $0 < b < 1$).

ASSUMPTION 4C—*Scalar Unobservable:* *Firms' intermediate input demand is given by*

$$(24) \qquad m_{it} = \widetilde{f}_t(k_{it}, l_{it}, \omega_{it}).$$

ASSUMPTION 5C—*Strict Monotonicity:* $\widetilde{f}_t(k_{it}, l_{it}, \omega_{it})$ *is strictly increasing in* ω_{it}.

First, note that we consider a "value-added" production function in the sense that the intermediate input m_{it} does not enter the production function to be estimated. One interpretation of this is that the gross output production function is Leontief in the intermediate input, where this intermediate input is proportional to output (e.g., see our Monte Carlo specification, or Gandhi, Navarro, and Rivers (2014)). Another follows the existing literature on value-added production functions; see, for example, Bruno (1978) and Diewert (1978). As discussed in those papers, as well as in Basu and Fernald (1997), the assumptions behind the existence of a meaningful value-added production function are subtle and nontrivial. We would not suggest applying our procedure to gross output production functions that are not Leontief in the intermediate input(s) (i.e., production functions where m_{it} enters (23)). This is because Bond and Söderbom (2005) (for the Cobb–Douglas case) and Gandhi, Navarro, and Rivers (2014) (for the general case) have shown that under the scalar unobservable assumptions of OP, LP, and our procedure, these gross output production functions are not identified without imposing further restrictions of the model. If one is interested in doing this, it is likely preferable to use an approach where OP/LP-like moments are augmented by first order conditions or revenue share equations, for example, LP (Appendix B) and Doraszelski and Jaumandreu (2013), who did this in the context of a Cobb–Douglas production function, or Gandhi, Navarro, and Rivers (2014), who show how to do this much more generally.

Moving to the assumptions of the model, Assumption 3c generalizes Assumption 3 by allowing labor to have dynamic implications, that is, choice of l_{it} not only affects current profits, but future profits, for example, through hiring or firing costs. Note that this implies that l_{it} is part of the state space of the firm's dynamic problem. We can allow l_{it} to be chosen at period t (and thus a function of I_{it}), chosen at period $t-1$ (and thus a function of only I_{it-1}), or at some point in between (with an intermediate information set).

Assumption 4c represents the crucial conceptual difference between our procedure and LP. The difference between Assumption 4b and Assumption 4c is that we are using a *conditional* input demand function to control for unobserved productivity, while LP uses an *unconditional* (or less conditional) input demand function. More precisely, (24) is the input demand function for m_{it} *conditional* on l_{it}, while (10) is the input demand function for m_{it} *not conditional* on l_{it}. While our conditional input demand function might be most naturally thought of as representing a situation where m_{it} is chosen *after* l_{it}, it can

equally well represent a situation where m_{it} and l_{it} are chosen simultaneously.[10] As does Assumptions 4 and 4b, Assumption 4c rules out production functions with multiple structural unobservables like McElroy (1987). Assumption 5c is the strict monotonicity assumption necessary for the inversion; since m_{it} is still a non-dynamic input, it is straightforward to apply the invertibility proof of LP to obtain this result.

Given these assumptions, we can follow LP, invert intermediate input demand $\omega_{it} = \widetilde{f}_t^{-1}(k_{it}, l_{it}, m_{it})$, and substitute into the production function, that is,

$$(25) \qquad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \widetilde{f}_t^{-1}(k_{it}, l_{it}, m_{it}) + \varepsilon_{it} = \widetilde{\Phi}_t(k_{it}, l_{it}, m_{it}) + \varepsilon_{it}.$$

Since we follow LP and treat $\widetilde{f}_t^{-1}$ nonparametrically, the first three terms are clearly not identified and are subsumed into $\widetilde{\Phi}_t(k_{it}, l_{it}, m_{it}) = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it}$, resulting in the following first stage moment condition:

$$(26) \qquad E[\varepsilon_{it} \mid I_{it}] = E\big[y_{it} - \widetilde{\Phi}_t(k_{it}, l_{it}, m_{it}) \mid I_{it}\big] = 0.$$

Unlike LP, (26) does not permit estimation of β_l in the first stage. However, it does produce an estimate $\widehat{\widetilde{\Phi}}_t(k_{it}, l_{it}, m_{it})$ of $\widetilde{\Phi}_t(k_{it}, l_{it}, m_{it})$. We propose estimating β_l along with the other production function parameters in the second stage using the following second stage conditional moment:

$$(27) \qquad E[\xi_{it} + \varepsilon_{it} \mid I_{it-1}]$$
$$= E\big[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it}$$
$$\quad - g\big(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}\big) \mid I_{it-1}\big] = 0,$$

where $\widetilde{\Phi}_{t-1}$ is replaced by its estimate from the first stage. Note that this moment condition is essentially identical to that used by LP (i.e., equation (14)), except for the fact that since $\widetilde{\Phi}_{t-1}$ now contains the labor term, it also needs to be subtracted out inside the g function.

As usual, it is easiest to transform conditional moments into unconditional moments for actual estimation. It is important to note that since our second stage requires estimating an additional parameter (β_l) as compared to LP, we require an additional unconditional moment relative to LP. To illustrate, consider a simple model where $\omega_{it} = \rho \omega_{it-1} + \xi_{it}$. Suppose the first stage is performed by OLS regression of y_{it} on a high-order polynomial in (k_{it}, l_{it}, m_{it}) to obtain $\widehat{\widetilde{\Phi}}_t(k_{it}, l_{it}, m_{it})$. If labor is assumed to be chosen after time $t-1$ (imply-

---

[10]There is no loss of generality in representing $\max_{l_{it}, m_{it}}\{\cdot\}$ as $\max_{l_{it}}\{\max_{m_{it}|l_{it}}\{\cdot\}\}$. Note that LP's unconditional (on l_{it}) intermediate input demand equation, that is, $m_{it} = f_t(k_{it}, \omega_{it})$, is most naturally interpreted as a situation where l_{it} and m_{it} are chosen simultaneously as a function of the current state (k_{it}, ω_{it}).

ing l_{it} will generally be correlated with ξ_{it}), a natural set of four second stage moment conditions to estimate the three production function parameters (β_0, β_k, and β_l) and ρ is

$$(28) \quad E\left[\left(y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it}\right.\right.$$
$$\left. - \rho \cdot \left(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}\right)\right)$$
$$\otimes \begin{pmatrix} 1 \\ k_{it} \\ l_{it-1} \\ \widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) \end{pmatrix}\right]$$
$$= 0.$$

In contrast, second stage LP estimation based on (14) only requires three moments, though as exemplified in the LP empirical work, one could alternatively utilize all four moments (or more than four using other components of I_{it-1}) and have overidentifying restrictions.[11]

In some industries, one might be willing to assume that labor is chosen by the firm at $t - 1$, that is, $l_{it} \in I_{it-1}$ (or alternatively make the assumption that ω_{it} is not observed by the firm until period $t + 1$). This is a potentially strong assumption, but it might be plausible in situations where there are significant hiring or firing costs, or labor market rigidities, possibly due to government regulation. In this case, one could add l_{it} to the information set (or replace l_{it-1} with l_{it} in the instrument set if one wanted to keep the model exactly identified), that is,

$$(29) \quad E\left[\left(y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it}\right.\right.$$
$$\left. - \rho \cdot \left(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}\right)\right.$$

---

[11]There are many ways to estimate the model based on the second stage moments. For example, the Appendix in LP illustrates how one can do second step estimation with a nonlinear search over just (β_k, β_l). A similar approach can work here, where one only has to search over (β_k, β_l) rather than over (β_0, β_k, β_l) and the parameters of g (see our Monte Carlo code). Relatedly, a prior version of this paper illustrated the approach with a second step moment in ξ_{it} rather in $\xi_{it} + \varepsilon_{it}$. As noted by Gandhi, Navarro, and Rivers (2014), the moment in $\xi_{it} + \varepsilon_{it}$ can be used more generally, for example, in some cases where the investment or intermediate input variable is only *weakly* monotonic in ω_{it}. Note that there are additional conditional moment conditions implied by the model related to inputs further in the past. Whether these are used as additional moments is typically a matter of preference.

$$\otimes \begin{pmatrix} 1 \\ k_{it} \\ l_{it} \\ l_{it-1} \\ \widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) \end{pmatrix} \Bigg] $$

$$= 0.$$

This stronger assumption will generally lead to more precise estimates.

4.2. *Discussion*

In the context of the intermediate input function approach, we see a number of advantages of using our *conditional* intermediate input demands over LP's *unconditional* intermediate input demands. First, because we do not attempt to estimate β_l in the first step, we avoid the functional dependence issues in the LP first stage. As a result, consistent estimates of β_l do not rely on DGPs that involve optimization error in l_{it} (and none in m_{it}), or i.i.d. firm-specific wage or output price shocks that are realized after the firm's choice of m_{it} (though our approach can provide consistent (but not as efficient) estimates of β_l under such DGPs). Second, as noted above, this model is consistent with labor being a dynamic input. With such dynamics, in the case where m_{it} and l_{it} are chosen simultaneously, Assumption 4b does not generally hold, since unconditional on l_{it}, m_{it} will depend on l_{it-1}. However, Assumption 4c does hold, since conditional on l_{it}, l_{it-1} should not affect a firm's optimal choice of m_{it}.

Third, estimation using our moments (26) and (27) produces consistent estimates under some DGPs where the canonical LP moments (13) and (14) do not. One important example is a situation in which there is across-firm variation in exogenous wage conditions (or adjustment costs to dynamic labor) that is potentially serially correlated over time. Suppose that this wage (or adjustment cost) variation is not observed by the econometrician. In this case, the *conditional* intermediate input demand function does not depend on the wage/adjustment cost, that is, conditional on l_{it}, a firm's choice of m_{it} *does not* depend on the wage/adjustment cost. On the other hand, firms' optimal choices of m_{it} given only k_{it}, that is, the *unconditional* intermediate input demand function, *do* generally depend on the unobserved wage/adjustment cost. In other words, Assumption 4c does hold, while Assumption 4b does not. Of course, it should be noted that if there are unobserved shocks to either prices of intermediate inputs or demand conditions, *neither* Assumption 4b or Assumption 4c holds, that is, neither the conditional or unconditional approach will produce consistent estimates. To summarize with respect to the above and the discussion in Section 2.1, consider serially correlated, exogenous, unobserved shocks to the costs of k_{it}, l_{it}, and m_{it}. OP cannot allow any of these shocks, LP can allow those relating to k_{it}, but not those relating to l_{it} and m_{it}, and our

proposed moments can allow those relating to k_{it} and l_{it}, but not those relating to m_{it}.

4.3. *Extensions*

4.3.1. *Investment Function Approach*

One can also use our methodology with the investment demand function of OP, replacing the unconditional investment demand function $i = f_t(k_{it}, \omega_{it})$ with a conditional (on l_{it}) investment demand function $i_{it} = f_t(k_{it}, l_{it}, \omega_{it})$. As above, this precludes identification of β_l in the first stage; again, in our view this is a benefit as it avoids the first stage functional dependence issues. After a first stage that estimates $y_{it} = \widetilde{\Phi}_t(k_{it}, l_{it}, i_{it}) + \varepsilon_{it}$, the second stage moments

$$
\begin{aligned}
&E[\xi_{it} + \varepsilon_{it} \mid I_{it-1}] \\
&= E\big[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} \\
&\quad - g\big(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, i_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}\big) \mid I_{it-1}\big]
\end{aligned}
$$

can be used to estimate the parameters. As when using the intermediate input demand function, this procedure is consistent with labor having dynamic effects. However, unlike when using the intermediate input demand function, it is not generally consistent with other unobservables entering either the i_{it} or l_{it} decisions. Any unobservable other than ω_{it} that directly impacts investment, for example, a firm-specific shock to the price or adjustment costs of capital, is clearly problematic. Less obviously, serially correlated unobservables that affect the l_{it} decision, for example, firm-specific wage shocks, will generally also affect the i_{it} decision directly since i_{it} is a dynamic decision variable. For example, a firm facing low wages today expects to face low wages in the future, and hence invests more today. As a result, the inversion is problematic. The reason the intermediate input demand function approach of LP is more robust to these additional serially correlated unobservables is because intermediate inputs have no dynamic aspects to them, that is, they only impact current output.

4.3.2. *Joint Estimation*

Wooldridge (2009) suggested an alternative implementation of the OP/LP moments that involves minimizing the first and second stage moments simultaneously. For example, in the LP model, he suggested estimating all the parameters simultaneously using the moment conditions

$$
\begin{aligned}
(30) \quad &E\begin{bmatrix} \varepsilon_{it} \mid I_{it} \\ \xi_{it} + \varepsilon_{it} \mid I_{it-1} \end{bmatrix} \\
&= E\begin{bmatrix} y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, m_{it}) \mid I_{it} \\ y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - g\big(\Phi_{t-1}(k_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1}\big) \mid I_{it-1} \end{bmatrix} \\
&= 0.
\end{aligned}
$$

As pointed out by Wooldridge, there are several advantages to this approach. First, the joint approach avoids the functional dependence issue described above. Even if l_{it} is functionally dependent on m_{it}, k_{it}, and t, β_l might be identified by the second set of moments (however, the nonparametric non-identification arguments of Gandhi, Navarro, and Rivers (2014) might apply to such an approach). Other advantages of the joint approach are potential efficiency gains, and simpler standard error calculations. There are also disadvantages of the joint approach; in particular, the joint approach requires a nonlinear search over β_0, β_k, β_l, *and* the parameters representing the two unknown functions Φ_t and g. This is more time consuming and likely more error prone than the two-stage approach, which can often be performed with a nonlinear search over just β_k and β_l (if Φ_t and g are represented by polynomials, the first stage can be estimated using OLS, and in the second stage, OLS can be used to concentrate out the parameters of g—see the Appendix).

Note that if one applied the idea of Wooldridge directly to our proposed moments (26) and (27), one gets something slightly different than (30), that is,

$$(31) \qquad E\begin{bmatrix} \varepsilon_{it} \mid I_{it} \\ \xi_{it} + \varepsilon_{it} \mid I_{it-1} \end{bmatrix}$$

$$= E\begin{bmatrix} y_{it} - \widetilde{\Phi}_t(k_{it}, l_{it}, m_{it}) \mid I_{it} \\ y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - g(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}) \mid I_{it-1} \end{bmatrix}$$

$$= 0.$$

Specifically, (30) is a restricted version of (31), a version where $\widetilde{\Phi}_t(k_{it}, l_{it}, m_{it})$ is assumed to be linear in l_{it}. The distinction also exactly corresponds to our distinction between moments based on the unconditional intermediate input demand function (the Wooldridge moments (30)) and our moments based on the conditional intermediate input demand function (31). For the same reasons as discussed above, we prefer (31) because it makes fewer assumptions—(31) allows serially correlated unobserved wage shocks, allows dynamic labor with adjustment costs, and allows materials to be chosen after labor (moreover, these potential aspects of the data generating process help avoid the non-identification results of Gandhi, Navarro, and Rivers (2014)). On the other hand, (30) does not allow this. One could also use the two sets of moments to test this restriction; if one models and interprets $\widetilde{\Phi}_t$ as a flexible parametric polynomial (rather than a nonparametric function), one could use a simple Wald or LR test for the restrictions that the linear coefficient on l in $\widetilde{\Phi}(k, l, m)$ equals β_l and all other coefficients involving l equal 0 (note that this is related to the overidentifying restrictions tests performed by OP and LP).

4.3.3. *Relation to Dynamic Panel Methods*

Interestingly, the form of our suggested estimators makes them fairly easy to compare to estimators used in an alternative literature, the dynamic panel literature. This is important because up to now, researchers interested in estimating

production functions have essentially been choosing between the OP/LP general approach versus the dynamic panel approach without a clear description of the similarities and differences of the identifying assumptions used in the two methods. We briefly discuss some differences and distinct advantages and disadvantages of both approaches.

As developed by work such as Chamberlain (1982), Anderson and Hsiao (1982), Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998, 2000), the dynamic panel (DP) literature essentially extends the fixed effects literature to allow for more sophisticated error structures. We consider a simplified version of these models that is closest to our model:

$$(32) \qquad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it},$$

where ω_{it} follows an AR(1) process, that is, $\omega_{it} = \rho \omega_{it-1} + \xi_{it}$. Like in our model, assume that ε_{it} is i.i.d. over time and uncorrelated with I_{it} and that ω_{it} is correlated with k_{it} and l_{it} $\forall t$, but that the innovation ξ_{it} is uncorrelated with I_{it-1}, that is, all input choices *prior* to t.

The dynamic panel literature might proceed by "ρ-differencing" this model, that is,

$$
\begin{aligned}
(33) \quad & y_{it} - \rho y_{it-1} \\
& = \beta_0(1 - \rho) + \beta_k(k_{it} - \rho k_{it-1}) \\
& \quad + \beta_l(l_{it} - \rho l_{it-1}) + \xi_{it} + (\varepsilon_{it} - \rho \varepsilon_{it-1}),
\end{aligned}
$$

and estimating the model using the moment condition

$$
\begin{aligned}
(34) \quad & E\big[\xi_{it} + (\varepsilon_{it} - \rho \varepsilon_{it-1}) \mid I_{it-1}\big] \\
& = E\big[(y_{it} - \rho y_{it-1}) - \beta_0(1 - \rho) \\
& \quad - \beta_k(k_{it} - \rho k_{it-1}) - \beta_l(l_{it} - \rho l_{it-1}) \mid I_{it-1}\big] \\
& = 0.
\end{aligned}
$$

This moment condition is similar to our proposed second stage moment condition $E[\xi_{it} + \varepsilon_{it} \mid I_{it-1}] = 0$. What is most relevant is the alternative assumptions under which they were derived. The DP approach does not need the assumptions that generate invertibility of the variable input demand function. So, for example, it can allow for unobserved cost shocks to *all* inputs, unlike our approach, which does not allow such shocks to the price of m_{it}. On the other hand, the DP derivation seems to rely on the linearity of the ω_{it} process; in contrast, OP, LP, and our approach can treat the first order Markov process completely nonparametrically. There are other differences between the models. For example, the DP literature can be extended to allow for a fixed effect α_i in addition to the AR(1) process, while generally speaking, this is challenging

in our context because it would tend to violate the scalar unobservable assumption. The DP literature also can potentially allow future values of the intermediate input or investment variable to depend on past ε_{it}'s, while our approach cannot. On the other hand, as elaborated on in OP, the scalar unobservable assumption of OP/LP and our approach makes it fairly straightforward to extend the methodologies to address endogenous exit (selection) from a sample (see OP); this would be considerably harder in the DP context. In summary, both approaches require strong (but different) assumptions. In some cases, a priori beliefs about a particular production process and/or data considerations may guide choices between the two approaches. In other cases, one may want to try both techniques. Finding that estimates are consistent across multiple techniques with different assumptions is surely more convincing than only using one.

4.3.4. *Units of Measure*

The above discussion has implicitly assumed an "ideal" data scenario where all the variables are measured in physical units that are equivalent across firms. While l_{it} is often measured in physical units (number of workers, or number of hours), output and inputs such as capital and materials are probably most frequently measured in monetary units. When output or an input is measured in monetary values, there are a few things that must be considered to apply the above techniques.

First, assume that firms are price takers in both input and output markets. The first question is whether the prices of equivalent inputs or output vary across firms.[12] If there is no variation in prices, then the monetary units are comparable across firms and one can apply the above techniques straightforwardly. On the other hand, if the prices of those equivalent inputs (or output) do vary across firms, then the monetary units are not comparable across firms (see, e.g., Klette and Griliches (1996)). In this case, the price variation must be observed for straightforward application of the above techniques. When the price variation is observed, then one can easily divide by prices to obtain measures of physical units and one can apply the above (note that the observed input and output prices need to be included in the intermediate input (or investment) demand function). Note that the observed input prices could also be used as additional instruments for the moment conditions under appropriate assumptions (e.g., that those input prices are exogenous); see, for example, Doraszelski and Jaumandreu (2013).

---

[12]When we describe measures of inputs and outputs as being equivalent across firms, we are ruling out a situation where input or output price differences across firms reflect differences in quality. If one thinks that quality differences are important, monetary measures of input or outputs may be superior to (non-equivalent) physical measures, though as noted below, using monetary measures requires an assumption that firms face the same menu of prices (menu over different quality levels).

Alternatively, consider a situation where firms face downward sloping output demand curves (and/or upward sloping input supply curves). In this case, one will generally need to assume that firms are all facing identical demand/supply curves; otherwise, firms will have different intermediate input (or investment) demand functions (i.e., the scalar unobservable Assumptions 4, 4b, or 4c will be violated). But even if this is true, there is still a problem since firms will generally be operating on different points of those demand and supply curves. Because of the resulting price differences, monetary values of inputs or output will again generally not be comparable across firms. If one observes the price differences of equivalent inputs or outputs, one can again construct measures of equivalent physical units and apply the above. If one does not observe the price differences, then other techniques are required—see Klette and Griliches (1996), De Loecker (2011), and De Loecker and Warzynski (2012) for recent methodological advances for this situation.

5. MONTE CARLO EXPERIMENTS

The discussion above implies that whether OP, LP, or our procedure generates consistent estimates of the production function parameters depends on the underlying DGP. What we instead focus on in our Monte Carlo experiments is how the various procedures perform under misspecification. Specifically, we consider misspecification where the intermediate input or investment variable is measured with error. In this case, none of the procedures produce consistent estimates, but we illustrate (in at least one setup) that our procedure appears less affected by this misspecification than OP and LP. Note that since the original version of the paper, others have proposed newer methods of explicitly addressing measurement error in inputs; Huang and Hu (2011) proposed a method to explicitly deal with such measurement error in these approaches based on observing multiple proxies and deconvolution methods, and Kim, Petrin, and Song (2013) also allowed for measurement error in capital.

Our Monte Carlo setup extends the work of Syverson (2001) and Van Biesebroeck (2007), who considered a simple dynamic model of firm investment and production that is analytically solvable (which considerably eases computation burden and reliability issues for Monte Carlo purposes). The parameters of the model were chosen to match a couple of key moments in the Chilean data used by LP. The full model setup is described in the Appendix, but the basics are as follows. Productivity follows a first order AR(1) process. Firms make optimal choices of investment in the capital stock to maximize the expected (discounted) value of future profits, where there are convex capital adjustment costs, and as above, the period-t capital stock is determined by investment at $t-1$ (i.e., $K_{it} = (1-\delta)K_{it-1} + I_{it-1}$). Material inputs m_{it} are chosen at t, while labor input l_{it} is chosen either at t or at $t-b$ (in the latter case, labor is chosen with only knowledge of ω_{it-b}, not ω_{it}). The production function is assumed

Leontief in (and proportional to) materials (as in Van Biesebroeck (2003), for example), that is,

$$Y_{it} = \min\{\beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} \exp(\omega_{it}), \beta_m M_{it}\} \exp(\varepsilon_{it}),$$

where $\beta_0 = 1$, $\beta_k = 0.4$, $\beta_l = 0.6$, and $\beta_m = 1$. Given that the Leontief first order condition holds, this means that we can proceed using the "structural value added" (see Gandhi, Navarro, and Rivers (2014)) (log) production function

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it}.$$

The other free parameters of the model were chosen such that the steady state of the model approximately matches the overall relationship between capital and labor in the Chilean data used by LP. Specifically, in our baseline specifications, 95% of the variation in capital in the data is across-firm variation, and the R^2 of a regression of capital on labor is approximately 0.5. Importantly, to match these two "moments" of the Chilean data, we needed to allow unobserved heterogeneity across firms in capital adjustment costs (see the Appendix for details). This violates the scalar unobservable assumption if one is using the investment variable of OP, and highlights the advantages of the LP intermediate input variable approach over the investment variable approach as described earlier; that is, the intermediate input approach (either with the LP moments or with our moments) allows unobserved variation across firms in investment prices or capital adjustment costs. As a result, we primarily compare results using LP with the intermediate input variable, to our approach also using the intermediate input variable.

We consider two base DGPs. The first, DGP1, is favorable to our procedure in that we add aspects to the DGP under which our moments produce consistent estimates but the LP moments do not. First, we assume that firms face different (unobserved to the econometrician) wages, where the wage process for firm i follows an AR(1) process. Second, we assume labor is chosen at time $t - b$, where $b = 0.5$ and where ω follows an AR(1) process both between ω_{it-1} and ω_{it-b} and between ω_{it-b} and ω_{it}. Details are in the Appendix. As discussed in Section 3, the reason that our moments produce consistent estimates in this DGP is because we use the conditional (on l_{it}) intermediate demand function, while LP does not.[13]

---

[13]Note that, along with the assumption that firms are price takers in input and output markets, this particular "proportional in M_{it}" Leontief specification avoids Gandhi, Navarro, and Rivers's (2014) critique of using these techniques with Leontief production functions. More specifically, in this setup, capital and labor will not be idled unless the firm chooses to shut down (if the scaled price of materials is somehow less than the price of output), and if the firm shuts down, it would presumably not be in the data set. Moreover, there is not even this type of optimal shut down in our Monte Carlos, since there is no across-firm variation in the price of materials or output. See the Appendix for more details.

The second DGP (DGP2) is favorable to LP. There is no wage variation across firms, and labor is chosen at t with full information on ω_{it}, so the model is consistent with the LP moments. We add simple optimization error in labor, where the labor variable is essentially the optimal level of labor plus i.i.d. noise. As noted above, this seems to be the simplest way to overcome the functional dependence issue in a way that is consistent with the LP assumptions (recall that choice of m_{it} cannot respond to the optimization error in l_{it}, so in this DGP the m_{it} that is observed needs to be "planned materials," i.e., the m_{it} chosen prior to the realization of optimization error in l_{it}). While our procedure also should produce consistent estimates with this DGP, we suspect that our estimates will not be as efficient as LP, since in this case, the LP moments can be seen as a restricted version of our moments (where that additional restriction is valid). In the context of our moments, the additional (valid) restriction is that $\widetilde{\Phi}_t$ is linear in l_{it} (with the coefficient β_l).

We also consider a third DGP (DGP3) that is a combination of the above two, that is, it includes optimization error in l_{it}, serially correlated wage variation across firms, and labor being chosen at time $t-b$. Neither the LP moments nor our moments produce consistent estimates under DGP3. While our procedure produces consistent estimates under DGP1 or DGP2 individually, it does not produce consistent estimates under the combination of them because in DGP2, the m_{it} that is observed depends on l_{it} *without* optimization error. Hence, the inverted materials demand function that is used, that is, a function of the observed l_{it} *with* optimization error, is not correct.[14,15]

In each of these three DGPs, we start with a base case in which there is no measurement error in the materials variable.[16] We then add increasing levels of such i.i.d. measurement error. Again, as soon as this measurement error is added, neither LP nor our procedure produces consistent estimates; our goal is

---

[14]This turns out to not matter in DGP2 because in DGP2, m_{it} can be written as a function of just k_{it} and ω_{it}, and thus including the incorrect l_{it} is irrelevant. As noted in the Appendix, our procedure would produce consistent estimates under a combination of DGP1 and DGP2 if DGP2 were altered such that the m_{it} that is observed optimally responds to the optimization error in l_{it}. We did not use this alternative DGP2, because under this alternative assumption, LP is not consistent.

[15]For comparison purposes, standard OLS estimates of the production function across these three DGPs produce labor coefficients between 0.85 and 0.95 and capital coefficients between 0.06 and 0.17.

[16]As pointed out in a prior version of this paper, there is a identification caveat using our suggested moments in all three of these DGPs. More specifically, there is a "global" identification issue in that the moments have expectation zero not only at the true parameters, but also at one other point on the boundary of the parameter space where $\widehat{\beta}_k = 0$, $\widehat{\beta}_l = \beta_l + \beta_k$, and the estimated AR(1) coefficient on ω equals the AR(1) coefficient on the wage process. One can easily calculate that at these alternative parameter values, the second stage moment equals the innovation in the wage process, which is orthogonal to k_{it} and l_{it-1}. This "spurious" minimum is a result of labor satisfying a static first order condition, and we suspect it would not occur were labor to have dynamic implications, nor when the alternative moments (29) are assumed. As such, we ignore this spurious minimum in our Monte Carlos.

TABLE I

MONTE CARLO RESULTS[a]

	ACF				LP			
	β_l		β_k		β_l		β_k	
Meas. Error	Coef.	Std. Dev.	Coef.	Std. Dev.	Coef.	Std. Dev.	Coef.	Std. Dev.
DGP1—Serially Correlated Wages and Labor Set at Time $t-b$								
0.0	0.600	0.009	0.399	0.015	0.000	0.005	1.121	0.028
0.1	0.596	0.009	0.428	0.015	0.417	0.009	0.668	0.019
0.2	0.602	0.010	0.427	0.015	0.579	0.008	0.488	0.015
0.5	0.629	0.010	0.405	0.015	0.754	0.007	0.291	0.012
DGP2—Optimization Error in Labor								
0.0	0.600	0.009	0.400	0.016	0.600	0.003	0.399	0.013
0.1	0.604	0.010	0.408	0.016	0.677	0.003	0.332	0.011
0.2	0.608	0.011	0.410	0.015	0.725	0.003	0.289	0.010
0.5	0.620	0.013	0.405	0.017	0.797	0.003	0.220	0.010
DGP3—Optimization Error in Labor and Serially Correlated Wages and Labor Set at Time $t-b$ (DGP1 plus DGP2)								
0.0	0.596	0.006	0.406	0.014	0.473	0.003	0.588	0.016
0.1	0.598	0.006	0.422	0.013	0.543	0.004	0.522	0.014
0.2	0.601	0.006	0.428	0.012	0.592	0.004	0.473	0.012
0.5	0.609	0.007	0.431	0.013	0.677	0.005	0.386	0.012

[a] 1000 replications. True values of β_l and β_k are 0.6 and 0.4, respectively. Standard deviations reported are of parameter estimates across the 1000 replications.

to see how sensitive the various procedures are to the additional measurement error. The top third of Table I contains results for DGP1; in the first row, there is zero measurement error in the material input variable, and in later rows, the level of this measurement error is increased. Measurement error of, for example, 0.1 indicates that the measurement error is of a magnitude that increases the variance of m_{it} by 10%.

The first row demonstrates that, as expected, even without measurement error the LP moments do not produce what appear to be consistent estimates. On the other hand, our moment conditions do appear to produce consistent estimates when there is no measurement error. Interestingly, as we change the level of measurement error in m_{it}, the LP estimate of the labor coefficient is considerably more sensitive, ranging from way below the true value of 0.6 with no measurement error, to considerably above the true value when there is lots of such error. Estimates using our moment conditions are again fairly stable, though there seems to be a small positive bias as the measurement error increases from 0.

The middle third of Table I contains results from DGP2. As noted above, with no measurement error, both our procedure and the LP procedure effectively find the true coefficients on labor and capital. Also as expected, the LP estimates are considerably more precise than those using our moments. How-

ever, as measurement error in materials is added, the LP labor coefficient increases and the LP capital coefficients decrease away from the true parameters. The capital parameter using our moments seems to be relatively unaffected by the measurement error in materials, while the labor coefficients do move somewhat. However, generally speaking, the LP coefficient estimates seem more affected by the measurement error than do the coefficient estimates based on our moments.

The last third of Table I contains results from DGP3. Again, neither procedure is consistent in this case, even without measurement error in m_{it}. But without the measurement error, the estimates based on our moments are closer to the truth than the LP estimates. Moreover, we again find that our estimates are much less sensitive than the LP estimates to measurement error. For β_l and β_k, our moments generate estimates that are at most only about 10% away from the truth. In contrast, the LP estimate of the labor coefficient, for example, varies between 0.47 and 0.68. These results suggesting that our moments are more robust to misspecification than the LP moments seem consistent with results in a working paper version of this paper. There, using the actual Chilean data, we found that while results using our moments were relatively robust across a variety of intermediate input proxies (materials, electricity, fuels), results using the LP moments were more variable.

Lastly, we tried the OP estimator using investment to recover unobserved productivity. As noted previously, given our use of firm-specific capital adjustment costs to reproduce observed patterns in the data, OP will *not* be consistent under any of these DGPs. In fact, the misspecification due to the firm-specific adjustment costs seems to dwarf any misspecification due to measurement error (now in the investment variable). As a result, regardless of the level of that measurement error or the specification, we obtain labor coefficients estimates of approximately 0.82 to 0.87 (Table II). This again highlights how the general idea of Levinsohn and Petrin to use the intermediate input to invert productivity (particularly in the context of our alternative moment conditions) may be preferable to the original OP idea of using investment.

6. CONCLUSIONS

We argue that some popular production function estimation techniques introduced by Olley and Pakes (1996) and Levinsohn and Petrin (2003) may suffer, at least conceptually, from a functional dependence problem. We propose an alternative, though quite related, estimator that avoids this functional dependence problem. The alternative estimator differs from OP and LP in that it inverts input demand functions that are *conditional* on choice of labor input (while OP/LP invert *unconditional* input demand functions). We show how this allows for more general underlying DGPs than does LP, similarly to how LP allowed for some more general DGPs than did OP. In comparison to Wooldridge's (2009) stacked version of LP's moments (again based on unconditional input demands), our moments based on conditional input demands

TABLE II

Monte Carlo Results[a]

Meas. Error	OP			
	β_l		β_k	
	Coef.	Std. Dev.	Coef.	Std. Dev.
DGP1—Serially Correlated Wages and Labor Set at Time $t - b$				
0.0	0.824	0.002	0.188	0.004
0.1	0.836	0.002	0.179	0.004
0.2	0.845	0.002	0.171	0.003
0.5	0.865	0.002	0.151	0.003

[a] 1000 replications. True values of β_l and β_k are 0.6 and 0.4, respectively. Standard deviations reported are of parameter estimates across the 1000 replications.

are strictly less restrictive. Again, this means that our estimator is strictly more general than that proposed by Wooldridge, for example, allowing unobserved heterogeneity in labor input prices across firms, allowing labor to have dynamic implications, and allowing labor to possibly be chosen prior to materials with a different information set. As might be expected, this generality and robustness in comparison to Wooldridge and LP does come at an efficiency cost; that is, when the assumptions behind the *unconditional* input demand function approach are in fact correct, estimates based on our *conditional* approach are less efficient. These findings are born out by Monte Carlo experiments that compare performance of the estimators under various DGPs that are sometimes consistent with the assumptions behind the estimators and sometimes not.

APPENDIX: Monte Carlo Setup

We now describe our Monte Carlo setup. We consider a panel of 1000 firms over 10 time periods. The parameters of the model are chosen such that the capital and labor series approximate key aspects of the Chilean data used in LP. Specifically, in our baseline specifications, 95% of the variation in capital in the data is across-firm variation (vs. within-firm variation), and the R^2 of a regression of capital on labor is approximately 0.5.

A.1. *Production Function and Productivity Shocks*

We use a production function that is Leontief in the material input, that is,

$$(35) \qquad Y_{it} = \min\{\beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it}}, \beta_m M_{it}\} e^{\varepsilon_{it}},$$

where $\beta_0 = 1$, $\beta_k = 0.4$, $\beta_l = 0.6$, and $\beta_m = 1$. Note that in this specification, material input is proportional to output. This is perhaps the most natural Leon-

tief specification, corresponding to a situation where a specific, fixed amount of the material input needs to be used to produce one unit of output (e.g., some specific amount of metal required to produce a can to hold food, some amount of wood required to produce an item). ε_{it} is mean 0 measurement error with standard deviation 0.1 that is distributed i.i.d. over firms and time. The productivity shock ω_{it} follows the AR(1) process

$$(36) \qquad \omega_{it} = \rho \omega_{it-1} + \xi_{it},$$

where $\rho = 0.7$. The variances of the normally distributed innovation ξ_{it} (σ_ξ^2) and the initial value ω_{i0} ($\sigma_{\omega_{i0}}^2$) are set such that the standard deviation of ω_{it} is constant over time and equal to 0.3.

A.2. *Choice of Labor and Material Inputs*

As described in Section 5, our specifications differ in how they model the DGP generating the labor (and material) inputs. In all specifications, we assume that labor and materials are *not* dynamic inputs, that is, a firm's choice of L_{it} and M_{it} has no implication for profitability after period t. As noted in the main text, our procedure is consistent if labor has dynamic implications. We did not allow this in our Monte Carlos; for speed and reliability purposes, we wanted a model where the firms' dynamic programming problem can be solved analytically. When the only dynamic input is capital, we can extend Syverson (2001) and Van Biesebroeck (2007) to obtain a model for which the optimal investment level can be solved analytically. An alternative would be to follow Bond and Söderbom (2005), allow dynamics (e.g., adjustment costs) in both capital and labor, and solve the model numerically.

We consider two main data generating processes. DGP1 allows there to be firm-specific (unobserved to the econometrician) wage shocks, and also allows L_{it} to be chosen prior to period t, that is, without full knowledge of ω_{it}. DGP1 is favorable to our procedure in that our conditional input demand approach produces consistent estimates, but the LP unconditional approach does not. In contrast, DGP2 does not allow firm-specific wage shocks, and L_{it} is chosen with knowledge of ω_{it}. However, DGP2 does add optimization error to L_{it} (while there is no optimization error in M_{it}). DGP2 is favorable to LP, as the LP procedure produces consistent estimates. While our procedure also produces consistent estimates, we expect them to have higher standard errors than those generated by LP.

More specifically, DGP1 first allows labor to be chosen at time period $t - b$ (we set $b = 0.5$). We can think of decomposing the AR(1) process (36) into two subprocesses. First, ω_{it-1} evolves to ω_{it-b}, at which point in time the firm chooses labor input (as a function of ω_{it-b}). Then, after L_{it} is chosen, ω_{it-b}

evolves to ω_{it}. We use the following model of the evolution of ω between sub-periods:

$$(37) \quad \begin{aligned} \omega_{it-b} &= \rho^{1-b}\omega_{it-1} + \xi_{it}^A, \\ \omega_{it} &= \rho^b \omega_{it-b} + \xi_{it}^B. \end{aligned}$$

Thus, when $b > 0$, firms have less than perfect information about ω_{it} when choosing L_{it}, and when b increases, this information decreases. Note that this is consistent with the AR(1) coefficient in (36) since $\rho^{1-b}\rho^b = \rho$. Also for consistency with (36), we impose that $\mathrm{Var}(\rho^b \xi_{it}^A + \xi_{it}^B) = \mathrm{Var}(\xi_{it})$ and that the variance of ξ_{it}^A is such that the variance of ω_{it-b} is constant over time. This defines $\mathrm{Var}(\xi_{it}^A) \equiv \sigma_{\xi^A}^2$ and $\mathrm{Var}(\xi_{it}^B) \equiv \sigma_{\xi^B}^2$, and intuitively implies that $\sigma_{\xi^A}^2$ is decreasing in b and $\sigma_{\xi^B}^2$ is increasing in b. In DGP1, firms also face different wages where the (ln) wage process for firm i follows an AR(1) process

$$\ln(W_{it}) = 0.3 \ln(W_{it-1}) + \xi_{it}^W,$$

where the variances of the normally distributed innovation ξ_{it}^W ($\sigma_{\xi^W}^2$) and the initial value $\ln(W_{i0})$ are set such that the standard deviation of $\ln(W_{it})$ is constant over time and equal to 0.1. Relative to a baseline in which all firms face the mean log wage in every period, this wage variation increases the within-firm, across-time, standard deviation of $\ln(L_{it})$ by about 10%.

Given DGP1, firms optimally choose L_{it} to maximize expected profits by setting (with the difference between the price of output and the price of the material input normalized to 1)

$$L_{it} = \beta_0^{1/(1-\beta_l)} \beta_l^{1/(1-\beta_l)} W_{it}^{-1/(1-\beta_l)} K_{it}^{\beta_k/(1-\beta_l)} e^{(1/(1-\beta_l))(\rho^b \omega_{it-b} + (1/2)\sigma_{\xi^B}^2)}.$$

Note that DGP1 is not subject to the critique of Gandhi, Navarro, and Rivers (2014) of Leontief production functions (as noted in footnote 13 of the main paper). They discussed how, in the context of a Leontief production function, for example,

$$(38) \quad Y_{it} = \min\{\beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it}}, m(M_{it})\} e^{\varepsilon_{it}},$$

it is possible that the Leontief "first order condition"

$$(39) \quad m(M_{it}) = \beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it}}$$

might not hold. The intuition for their argument is as follows. Suppose K_{it} and L_{it} are chosen prior to M_{it} being chosen. Suppose that between these two points in time, the price of M_{it} rises precipitously (and unexpectedly), such that the revenue from selling output does not even cover the cost of the material input required to produce that output (i.e., ignoring the "sunk" costs of already

purchased K_{it} and L_{it}). In this case, firms would not generally choose M_{it} to satisfy (39), and thus the data could contain points where

$$Y_{it} \neq \beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it} + \varepsilon_{it}},$$

invalidating all the above procedures. However, in our Monte Carlo setup where output is proportional to material input (i.e., $m(M_{it}) = \beta_m M_{it}$, which could be the most plausible Leontief specification for many material inputs), this is not a problem. This is because firms will either satisfy (39) or produce zero output (the latter if the price of M_{it} unexpectedly rises enough after the other inputs are chosen), and if they produce zero, they will presumably not be in the data set of firms that are operating and thus not be a problem for estimation. Moreover, in our Monte Carlos, the prices of material inputs (and output) are constant across firms anyway, so firms never want to let capital and labor go unutilized and produce zero.[17]

In contrast to DGP1, in DGP2 firms face *identical* wages and L_{it} is chosen *at period t* (i.e., $b = 0$). These changes are needed to make the model consistent with the assumptions of LP. Moreover, we need something to break the functional dependence issue, so we add optimization error to L_{it}. More specifically, let L_{it}^{err} be the firm's choice of labor including optimization error. We assume that

$$L_{it}^{\text{err}} = L_{it} e^{\xi_{it}^l},$$

where $\xi_{it}^l \sim N(0, \sigma_{\xi^l}^2)$ is optimization error that is distributed i.i.d. over firms and time (and independent of everything else in the model). We set $\sigma_{\xi^l} = 0.37$, which also increases the within-firm, across-time, standard deviation of $\ln(L_{it})$ by about 10%.

Constructing this DGP such that LP produces consistent estimates is somewhat subtle. If the firm observes L_{it}^{err} prior to choosing the material input, they will typically choose

$$M_{it}^{\text{err}} = \beta_0 K_{it}^{\beta_k} L_{it}^{\text{err}\beta_l} e^{\omega_{it}}.$$

This is a problem if the econometrician observes M_{it}^{err} as the material input measure, as this would violate the scalar unobservable assumption of LP (Assumption 4b). So instead we assume that the econometrician observes M_{it} instead of M_{it}^{err}, where

$$M_{it} = \beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it}}.$$

---

[17]Note that in this example, a fully nonparametric first stage regression of Y_{it} on K_{it}, L_{it}, and M_{it} to recover ε_{it} is not strictly needed given the assumed Leontief functional form. For example, one could alternatively regress Y_{it} on just M_{it} to recover ε_{it}, or similarly, use the Leontief FOC directly. In our Monte Carlos, we do the fully general first stage (perhaps sacrificing some efficiency) because that corresponds to our more general procedure.

One interpretation of this is that, after observing ω_{it}, firms commit to "planned" optimal L_{it} and M_{it}. However, workers are randomly sick, and when more workers are sick less materials are used, so that the firm ends up using L_{it}^{err} and M_{it}^{err} and producing[18]

$$Y_{it} = \beta_0 K_{it}^{\beta_k} L_{it}^{\mathrm{err}\,\beta_l} e^{\omega_{it}} e^{\varepsilon_{it}}.$$

What the econometrician is assumed to observe are Y_{it}, K_{it}, L_{it}^{err}, and M_{it}. The interpretation of observing M_{it} is that what is reported in the data is planned (or ordered) materials, that is, prior to the realization of sick days and L_{it}^{err}.

In DGP2, both LP and our procedure produce consistent estimates. While we do not present results for this, one can consider an alternative DGP2, call it DGP2b, where M_{it}^{err} is observed instead of M_{it}. In DGP2b, LP does not produce consistent estimates, but our suggested procedure does (since conditional on L_{it}^{err}, the scalar unobservable assumption is not violated). We also consider the combination of DGP1 and DGP2, that is, a DGP that includes across-firm wage variation, labor chosen at $t - b$, and optimization error in L_{it}. In this setup, neither LP nor our procedure produces consistent estimates (see the main text for a discussion of why our procedure does not produce consistent estimates under DGP1 + DGP2, even though it provides consistent estimates under DGP1 or DGP2 separately). While we again do not present results, our procedure would produce consistent estimates under DGP1 + DGP2b, while LP would not.

For all three DGPs (DGP1, DGP2, and DGP3 = DGP1 + DGP2), we additionally add various levels of measurement error to the observed material input. As noted in the main text, this renders both LP and our approach inconsistent. However, we find that estimation based on the LP moments appears to be more sensitive to this measurement error than does estimation based on our moments.

A.3. *Investment Choice and Steady State*

In contrast to non-dynamic labor and material inputs, we assume that capital is a dynamic input. Specifically, capital is accumulated through investment according to

$$K_{it} = (1 - \delta)K_{it-1} + I_{it-1},$$

---

[18]In theory, the Leontief critique of GNR could apply here. For example, if between the point in time in which L_{it} is chosen and L_{it}^{err} is realized, there is a large enough unexpected positive shock to the price of the material input, it is possible that if $L_{it}^{\mathrm{err}} > L_{it}$, the firm would not want to increase material input levels over M_{it}, that is, the unexpected supply of labor would optimally go unutilized. However, in our setup, there are no such shocks to the price of material input (or output), so everything is internally consistent. Note also that when the econometrician observes planned materials, as in this DGP, one cannot simply estimate the model directly off the Leontief first order condition.

where $(1 - \delta) = 0.8$, and investment is subject to convex adjustment costs given by

$$c_i(I_{it}) = \frac{\phi_i}{2} I_{it}^2,$$

where $1/\phi_i$ is distributed lognormally across firms (but constant over time) with standard deviation 0.6. While the existence of this unobserved heterogeneity in adjustment costs is consistent with the LP procedure and our procedure, it is not consistent with the scalar unobservable assumption in the OP context. This renders OP estimates inconsistent in all our Monte Carlos. We included it because, without this heterogeneity, it was extremely hard (in steady state, see below) to match the within-firm versus across-firm variation in capital and the correlation between capital and labor of the Chilean data sets as discussed above.[19]

As described by Syverson (2001) and Van Biesebroeck (2007), under the assumption of constant returns to scale, a pared-down version of the above can be solved analytically using Euler equation techniques. We have verified that their analytic result can be generalized to allow for the extensions in the above model, that is, persistent across-firm variation in wages, labor being chosen at $t - b$, and optimization error in L_{it}. Specifically, a Euler equation approach implies the following optimal investment rule (where β is the discount factor, set to 0.95 in our Monte Carlos):

$$
\begin{aligned}
I_{it} = {}& \frac{\beta}{\phi_i} \sum_{\tau=0}^{\infty} (\beta(1-\delta))^{\tau} \left(\frac{\beta_k}{1-\beta_l} \right) \beta_0^{1/(1-\beta_l)} \\
& \times \left[\beta_l^{\beta_l/(1-\beta_l)} e^{(1/2)\beta_l^2 \sigma_{\xi l}^2} - \beta_l^{1/(1-\beta_l)} e^{(1/2)\sigma_{\xi l}^2} \right] \\
& \times \exp\Bigg\{ \Bigg[\left(\frac{1}{1-\beta_l} \right) \rho^{\tau+1} \omega_{it} + \frac{-\beta_l}{1-\beta_l} \rho_W^{\tau+1} \ln(W_{it}) \\
& + \frac{1}{2} \left(\frac{-\beta_l}{1-\beta_l} \right)^2 \sigma_{\xi W}^2 \sum_{s=0}^{\tau} \rho_W^{2(\tau-s)} + \frac{1}{2} \left(\frac{1}{1-\beta_l} \right)^2 \rho_b^2 \Bigg(\rho^{2\tau} \sigma_{\xi a}^2 \\
& + \sum_{s=1}^{\tau} \rho^{2(\tau-s)} \sigma_{\xi}^2 \Bigg) + \left(\frac{1}{1-\beta_l} \right) \left(\frac{1}{2} \sigma_{\xi b}^2 \right) \Bigg] \Bigg\}.
\end{aligned}
$$

[19]Without heterogeneity in ϕ_i, the only way our DGPs can match 95% of the variation in K_{it} being "across-firm variation" is by setting the persistence of the productivity process ω_{it} extremely high. But when this persistence is extremely high, it implies a very high correlation between K_{it} and L_{it} which contradicts the R^2 of a fixed effects regression of L_{it} on K_{it} of only about 50% in the Chilean data, unless one includes unrealistic amounts of optimization error or wage variation.

Note that because of the constant returns to scale, optimal investment does not depend on current capital stock—the only firm-specific variables that it depends on are ω_{it}, ϕ_i, and $\ln(W_{it})$ (the latter because it is serially correlated over time).

For our experiments, we did not want our results to depend on an arbitrary, exogenous, initial distribution of capital stock across firms (one would also need to somehow specify correlations between K_{i0} and $(\omega_{i0}, \phi_i, \ln(W_{i0}))$). Instead, we decided to take data from the steady state distribution of firms implied by our model. To do this, we simply started all firms with $K_{i0} = 0$, and simulated firms into the future. Our 10 periods of data are taken from a point in time where the impact of the initial values appeared minimal, that is, the steady state.

A.4. *Estimation*

We estimate the following Leontief-derived value-added production function:

$$Y_{it} = \beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it}} e^{\varepsilon_{it}}$$

(or in DGP2, the same with L_{it} replaced with L_{it}^{err}). When we use our procedure or LP, we use M_{it} to invert the productivity shock, and when we run OP, we use I_{it}.[20]

For our procedure, our first stage is simply OLS of y_{it} on k_{it}, l_{it}, and m_{it}. More generally, one would want higher order polynomials in the explanatory variables, but because, in our setup, optimal m_{it} is actually a linear function of k_{it}, l_{it}, and ω_{it}, we simply used a first order polynomial. In the second stage, our estimation is based on the following moments (where $\widetilde{\beta}_0 = \ln(\beta_0)$):

$$(40) \quad E\left[\left(y_{it} - \widetilde{\beta}_0 - \beta_k k_{it} - \beta_l l_{it}\right.\right.$$

$$\left. - \rho \cdot \left(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \widetilde{\beta}_0 - \beta_k k_{it-1} - \beta_l l_{it-1}\right)\right)$$

$$\left. \otimes \begin{pmatrix} 1 \\ k_{it} \\ l_{it-1} \\ \widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) \end{pmatrix}\right]$$

$$= 0.$$

[20]As noted, one could also use our moments along with I_{it} as the variable to recover the productivity shock, but because of the unobserved heterogeneity in ϕ_i in our Monte Carlo setup, this would, like OP, not generate consistent estimates.

However, to reduce the dimension of the nonlinear search (and presumably increase speed and reliability for our experiments), we do not use these moments directly. Instead we "concentrate-out" two of the parameters $\widetilde{\beta}_0$ and ρ. More specifically, for a hypothetical guess of the parameters β_k and β_l, we construct

$$\widetilde{\beta_0 + \omega_{it}}(\beta_k, \beta_l) = \widehat{\widetilde{\Phi}}_t(k_{it}, l_{it}, m_{it}) - \beta_k k_{it} - \beta_l l_{it},$$

where, again, $\widehat{\widetilde{\Phi}}_t(k_{it}, l_{it}, m_{it})$ is the predicted value from the first stage. We then regress $\widetilde{\beta_0 + \omega_{it}}(\beta_k, \beta_l)$ on $\widetilde{\beta_0 + \omega_{it-1}}(\beta_k, \beta_l)$, noting that the residuals of this regression are the implied values of the innovations in omega, that is, $\widehat{\xi}_{it}(\beta_k, \beta_l)$. Note that this regression implicitly makes these innovations mean zero and uncorrelated with $\widehat{\omega_{it-1}}(\beta_k, \beta_l)$, so it is similar to enforcing the first and fourth moments in (40). While generally one might want to do this regression using higher order polynomials in the explanatory variable, we again simply do it linearly, since our true DGPs use a linear AR(1) process for ω_{it}. We then search over β_k and β_l space to find $\widehat{\beta}_k$ and $\widehat{\beta}_l$ that minimize the following moment conditions:

$$(41) \qquad E\left[\widehat{\xi}_{it}(\beta_k, \beta_l) \otimes \begin{pmatrix} k_{it} \\ l_{it-1} \end{pmatrix}\right] = 0.$$

Our estimate of ρ is given by the slope coefficient of the regression of $\widetilde{\beta_0 + \omega_{it}}(\beta_k, \beta_l)$ on $\widetilde{\beta_0 + \omega_{it-1}}(\beta_k, \beta_l)$ evaluated at $\widehat{\beta}_k$ and $\widehat{\beta}_l$.

For our LP estimates, we use the canned STATA routine (Petrin, Poi, and Levinsohn (2004))), edited to again use linear approximations (since according to the DGP, the true functions were in fact linear). Our results were very similar when we used the default routine based on cubic approximations. For our OP results, we also use the canned LP STATA routine modified to use the investment variable, but here we use the default cubic approximation, since the dynamic investment function is not linear. In the online supplemental materials, we include GAUSS versions of these estimation routines (Ackerberg, Caves, and Frazer (2015)).

REFERENCES

ACKERBERG, D., L. BENKARD, S. BERRY, AND A. PAKES (2007): "Econometric Tools for Analyzing Market Outcomes," in *Handbook of Econometrics*. Amsterdam: North-Holland. [2412]

ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2015): "Supplement to 'Identification Properties of Recent Production Function Estimators'," *Econometrica Supplemental Material*, 83, http://dx.doi.org/10.3982/ECTA13408. [2448]

ACKERBERG, D., X. CHEN, J. HAHN, AND Z. LIAO (2014): "Asymptotic Efficiency of Semiparametric Two-Step GMM," *Review of Economic Studies*, 81 (3), 919–943. [2419]

AI, C., AND X. CHEN (2003): "Efficient Estimation of Models With Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71 (6), 1795–1843. [2419,2423]

——— (2007): "Estimation of Possibly Misspecified Semiparametric Conditional Moment Restriction Models With Different Conditioning Variables," *Journal of Econometrics*, 141, 5–43. [2419]

ALVAREZ, R., AND R. LOPEZ (2005): "Exporting and Performance: Evidence From Chilean Plants," *Canadian Journal of Economics*, 38 (4), 1384–1400. [2412]

ANDERSON, T. W., AND C. HSIAO (1982): "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 18, 47–82. [2414,2434]

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. [2414,2434]

ARELLANO, M., AND O. BOVER (1995): "Another Look at the Instrumental Variable Estimation of Error Components Models," *Journal of Econometrics*, 68, 29–51. [2414,2434]

BASU, S., AND J. FERNALD (1997): "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 105, 249–283. [2428]

BLALOCK, G., AND P. GERTLER (2004): "Learning From Exporting: Revisited in a Less Developed Setting," *Journal of Development Economics*, 75 (2), 397–416. [2412]

BLUNDELL, R., AND S. BOND (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115–143. [2414,2434]

——— (2000): "GMM Estimation With Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19, 321–340. [2414,2434]

BOND, S., AND M. SÖDERBOM (2005): "Adjustment Costs and the Identification of Cobb–Douglas Production Functions," Working Paper Series No. 05/04, The Institute for Fiscal Studies. [2428, 2442]

BRUNO, M. (1978): "Duality, Intermediate Inputs, and Value-Added," in *Production Economics: A Dual Approach to Theory and Applications*, Vol. 2, ed. by M. Fuss and D. McFadden. Amsterdam: North-Holland, Chapter 1. [2428]

CHAMBERLAIN, G. (1982): "Multivariate Regression Models for Panel Data," *Journal of Econometrics*, 18, 5–46. [2414,2434]

CHEN, X., J. HAHN, Z. LIAO, AND G. RIDDER, (2013): "Asymptotic Properties of Nonparametric Two-Step Sieve Estimates," Report, UCLA. [2419]

CHRISTENSEN, L. R., AND W. H. GREENE (1976): "Economies of Scale in U.S. Electric Power Generation," *Journal of Political Economy*, 84 (4), 655–676. [2411,2415]

CHRISTENSEN, L. R., D. W. JORGENSON, AND L. J. LAU (1973): "Transcendental Logarithmic Production Frontiers," *Review of Economics and Statistics*, 55 (1), 28–45. [2411,2415]

CRISCUOLO, C., AND R. MARTIN (2009): "Multinationals and U.S. Productivity Leadership: Evidence From Great Britain," *Review of Economics and Statistics*, 91 (2), 263–281. [2412]

DE LOECKER, J. (2011): "Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 79 (5), 1407–1451. [2436]

DE LOECKER, J., AND F. WARZYNSKI (2012): "Markups and Firm-Level Export Status," *American Economic Review*, 102 (6), 2437–2471. [2436]

DIEWERT, W. E. (1978): "Hick's Aggregation Theorem and the Existence of a Real Value Added Function," in *Production Economics: A Dual Approach to Theory and Practice*, Vol. 2, ed. by M. Fuss and D. McFadden. Amsterdam: North-Holland, Chapter 1. [2428]

DOLLAR, D., M. HALLWARD-DREIMEIER, AND T. MENGISTAE (2005): "Investment Climate and Firm Performance in Developing Countries," *Economic Development and Cultural Change*, 54 (1), 1–31. [2412]

DORASZELSKI, U., AND J. JAUMANDREU (2013): "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies*, 80, 1338–1383. [2418,2422,2428,2435]

FERNANDES, A. (2007): "Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries," *Journal of International Economics*, 71 (1), 52–71. [2412]

GANDHI, A., S. NAVARRO, AND D. RIVERS (2014): "On the Identification of Production Functions: How Heterogeneous Is Productivity," Report, U–Wisconsin–Madison. [2428,2430,2433, 2437,2443]

GRILICHES, Z., AND J. HAUSMAN (1986): "Errors in Variables in Panel Data," *Journal of Econometrics*, 31, 93–118. [2414]

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. by S. Strøm. Cambridge, U.K.: Cambridge University Press. [2413-2415,2417, 2421]

GRILICHES, Z., AND V. RINGSTAD (1971): *Economies of Scale and the Form of the Production Function*. Amsterdam: North-Holland. [2414]

HALL, R. E. (1988): "The Relation Between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96, 921–947. [2414]

HOCH, I. (1955): "Estimation of Production Function Parameters and Testing for Efficiency," *Econometrica*, 23, 325–326. [2411,2414]

———— (1958): "Simultaneous Equation Bias in the Context of the Cobb–Douglas Production Function," *Econometrica*, 26, 566–578. [2411]

———— (1962): "Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data," *Econometrica*, 30, 34–53. [2411,2414]

HUANG, G., AND Y. HU (2011): "Estimating Production Functions With Robustness Against Errors in the Proxy Variables," Report, Johns Hopkins University. [2436]

KASAHARA, H., AND J. RODRIQUE (2008): "Does the Use of Imported Intermediates Increase Productivity? Plant Level Evidence," *Journal of Development Economics*, 87 (1), 106–118. [2412]

KIM, K., A. PETRIN, AND S. SONG (2013): "Estimating Production Functions With Control Functions When Capital Is Measured With Error," Report, University of Chicago. [2436]

KLEIN, L. R. (1953): *A Textbook of Econometrics*. New York: Prentice-Hall. [2414]

KLETTE, T. J., AND Z. GRILICHES (1996): "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous," *Journal of Applied Econometrics*, 11 (4), 343–361. [2435,2436]

LEVINSOHN, J., AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317–342. [2411,2413,2440]

MARSCHAK, J., AND W. H. ANDREWS JR. (1944): "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12, 143–205. [2411,2414]

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. [2411,2415,2417, 2418,2429]

MUNDLAK, Y. (1961): "Empirical Production Function Free of Management Bias," *Journal of Farm Economics*, 43, 44–56. [2411,2414]

———— (1963): "Estimation of Production and Behavioral Functions From a Combination of Cross-Section and Time-Series Data," in *Measurement in Economics*, ed. by C. Christ et al. Stanford, CA: Stanford University Press, 138–166. [2411,2414]

———— (1996): "Production Function Estimation: Reviving the Primal," *Econometrica*, 64 (2), 431–438. [2411,2424]

MUNDLAK, Y., AND I. HOCH (1965): "Consequences of Alternative Specifications in Estimation of Cobb–Douglas Production Functions," *Econometrica*, 33, 814–828. [2411,2414]

NERLOVE, M. (1963): "Returns to Scale in Electricity Supply," in *Measurement in Economics*, ed. by C. Christ et al. Stanford: Stanford University Press, 167–200. [2414]

OLLEY, S., AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1295. [2411,2413,2440]

OZLER, S., AND K. YILMAZ (2009): "Productivity Response to Reduction in Trade Barriers: Evidence From Turkish Manufacturing Plants," *Review of World Economics*, 145 (2), 339–360. [2412]

PAKES, A., AND S. OLLEY (1995): "A Limit Theorem for a Smooth Class of Semiparametric Estimators," *Journal of Econometrics*, 65, 295–332. [2421]

PANZAR, J. (1989): "Technological Determinants of Firm and Industry Structure," in *Handbook of Industrial Organization*, Vol. 1, ed. by R. Schmalensee and R. Willig. Amsterdam: North-Holland, 3–59. [2411]

PAVCNIK, N. (2002): "Trade Liberalization Exit and Productivity Improvements: Evidence From Chilean Plants," *Review of Economic Studies*, 69, 245–276. [2411,2412]

PETRIN, A., B. POI, AND J. LEVINSOHN (2004): "Production Function Estimation in Stata Using Inputs to Control for Unobservables," *The Stata Journal*, 4, 113–123. [2448]

ROBINSON, P. M. (1988): "Root-N-Consistent Semiparametric Regression," *Econometrica*, 56, 931–954. [2423]

SIVADASAN, J. (2009): "Barriers to Competition and Productivity: Evidence From India," *B.E. Journal of Economic Analysis & Policy*, 9 (1), Article 42. [2412]

SLADE, M. (1989): "Modeling Stochastic and Cyclical Components of Technical Change: An Application of the Kalman Filter," *Journal of Econometrics*, 41, 363–383. [2411]

SOLOW, R. M. (1957): "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39, 312–320. [2414]

SYVERSON, C. (2001): "Market Structuure and Productivity," Ph.D. Dissertation, University of Maryland. [2436,2442,2446]

TOPALOVA, P., AND A. KHADEWAL (2011): "Trade Liberalization and Firm Productivity: The Case of India," *Review of Economics and Statistics*, 93 (3), 995–1009. [2412]

VAN BIESEBROECK, J. (2003): "Productivity Dynamics With Technology Choice: An Application to Automobile Assembly," *Review of Economic Studies*, 70 (1), 167–198. [2437]

——— (2007): "Robustness of Productivity Estimates," *The Journal of Industrial Economics*, 55, 529–569. [2436,2442,2446]

WOOLDRIDGE, J. (2009): "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters*, 104 (3), 112–114. [2413,2419,2432, 2440]

Dept. of Economics, University of Michigan, Ann Arbor, MI 48109, U.S.A.; ackerber@umich.edu,

Economists Inc., Washington, DC 20037, U.S.A.; caves.k@east.ei.com,
and
Dept. of Management, University of Toronto Scarborough, Toronto, ON M1C 1A4, Canada, Rotman School of Management, University of Toronto, Toronto, ON M5S 3E6, Canada, and School of Public Policy and Governance, University of Toronto, Toronto, ON M5S 3K9, Canada; garth.frazer@rotman.utoronto.ca.