Production Function Estimation with Imperfect Proxies

Mert Demirer

January 11, 2023

Abstract

The 'proxy variable' approach is often used to estimate production functions. This approach is not robust to measurement error, and it relies on some strong assumptions, including strict monotonicity, scalar productivity, and timing. In this paper, I develop partial identification results that are robust to deviations from these assumptions and measurement errors in inputs. In particular, my model (i) allows for multi-dimensional unobserved heterogeneity, (ii) relaxes strict monotonicity to weak monotonicity, and (iii) accommodates a more flexible timing assumption for capital. I show that under these assumptions, production function parameters are partially identified by an 'imperfect proxy' variable via moment inequalities. Using these moment inequalities, I derive bounds on the parameters and propose an estimator.

MIT Sloan School of Management, email: mdemirer@mit.edu. This paper has benefited from feedback and suggestions from Anna Mikusheva, Nikhil Agarwal, Kirill Evdokimov, Jerry Hausman, Whitney Newey, Alex Olssen, Kamil Yilmaz, Nathan Zorzi as well as seminar participants at the MIT Econometrics lunch.

1 Introduction

Production functions are a critical input in many economic studies. These studies typically require estimating a production function using firm-level data. A major challenge in production function estimation is the endogeneity of inputs. Firms observe their productivity before choosing production inputs; however, productivity is unobservable to the researcher. A commonly used method to address the endogeneity problem is the proxy variable approach. Introduced by Olley and Pakes (1996) (henceforth OP), this approach relies on using a variable, called proxy, to control for unobserved productivity. OP use investment as a proxy variable, which is assumed to be a strictly increasing function of productivity conditional on capital. By inverting this unknown function, they essentially recover the productivity shock and control for it in the estimation. The proxy variable approach has become the workhorse for estimating production functions and has been extended by several papers. Levinsohn and Petrin (2003) (LP) have proposed using materials as a proxy, and Ackerberg et al. (2015) (ACF) have introduced a unified framework of proxy variable approach that deals with some practical concerns.

A limitation of the proxy variable approach is that it relies on strong assumptions, such as single-dimensional unobserved heterogeneity and strict monotonicity. These assumptions have important economic implications (Ackerberg et al. (2007), Ackerberg et al. (2015)). First, firms are differentiated only by a single productivity shock, which restricts firm-level heterogeneity. Second, there is no heterogeneity in adjustment costs and investment prices, as investment depends only on productivity. Third, estimation requires restricting competition in the output market. Moreover, the proxy variable approach is not robust to measurement errors in inputs, an important concern, especially for capital.

In this paper, I develop a partial identification approach that is robust to some proxy variable assumptions and measurement errors in inputs. In particular, my model (i) allows for multi-dimensional unobserved heterogeneity, (ii) relaxes strict monotonicity to weak monotonicity, (iii) accommodates a more general timing assumption, and (iv) is robust to measurement errors in all inputs. In my approach, the standard proxy variable becomes an 'imperfect proxy,' which can be used to derive moment inequalities for identification. Using these moment inequalities, I characterize the identified set for the parameters and propose an estimator.

An 'imperfect proxy' variable contains information about productivity, but it is not strictly monotone with it, so it cannot be directly used to control for productivity.

Instead, an imperfect proxy gives a stochastic ordering of productivity distributions, which can be used for identification. To show this result, I first group firms into 'high' investment firms, firms that invest more than a cutoff value, and 'low' investment firms, firms that invest less than the cutoff value. Then, I show that the productivity distribution of 'high' investment firms first-order stochastically dominates the productivity distribution of 'low' investment firms. The main idea for identification is to use this stochastic ordering in the form of moment inequalities to obtain bounds for production function parameters.

I derive moment inequalities and study identification under a wide range of assumptions. The first identification result relies only on the assumption that productivity shocks follow an exogenous Markov process. This is the least restrictive specification, and therefore, gives the widest bounds. The other identification results exploit modeling assumptions fully to derive moment inequalities via the imperfect proxy. These moment inequalities give an identified set for the main specification of the paper. I also show how to tighten the identified set under additional distributional assumptions and shape restrictions. Analyzing identification under a wide range of assumptions makes the role of each assumption in identification transparent. For example, one can start with the most general model to impose as few restrictions as possible. If the estimated set is not informative, then a nested model can be considered to shrink the identified set. Also, comparing the results from a nested model and a general model tests the restrictions imposed by the nested model.

The partial identification approach allows me to have a model with rich heterogeneity. My model includes two productivity shocks, one persistent and the other transitory. The firm can observe both of these shocks, so both can create endogeneity. Moreover, my model includes unobserved variables that affect the firm's choice of investment. Consequently, it allows for heterogeneity in input prices and adjustment costs as well as demand shock in the output market. Finally, the identification approach is robust to measurement errors in all inputs. This robustness is particularly crucial for capital, which is most prone to measurement error.

My method is generic in that it applies to production functions under different specifications. First, one can use my method to partially identify the parameters of both value-added or gross Cobb-Douglas production functions. Second, the model is agnostic about which proxy variable to use, so both investment and materials can serve as an imperfect proxy for estimation. Third, the model can accommodate different timing assumptions about capital. One can assume that capital is chosen one period in advance, as in prior approaches, or that firms choose capital after (partially) observing productivity shocks. Finally, the model is not specific to the Cobb-Douglas production function. A nonlinear production function that is known up to a finite-dimensional parameter vector can be considered.

1.1 Related Literature

This paper contributes to the large literature on production function estimation using proxy variables (OP, LP, ACF; Gandhi et al. (2020) (GNR)).[1] OP find the conditions under which investment can be used as a 'proxy' to control for unobserved productivity. Motivated by 'zero' and 'lumpy' investment problem, LP propose using materials as a proxy variable. ACF point out a collinearity issue in these papers and propose an alternative proxy variable approach that avoids the collinearity problem. My paper extends these approaches by showing how to make inferences when the standard proxy variable approach assumptions are violated.

A few recent papers study production function estimation with measurement errors in capital (Hu et al. (2011), Collard-Wexler and De Loecker (2016), Kim et al. (2016)). These papers require either an instrumental or another proxy variable to address measurement errors. In contrast, my method does not require an additional variable, but it gives a bound rather than a point estimate.

This paper is related to the literature on monotone instrumental variables (Manski (1997), Manski and Pepper (2000)). This literature assumes that the means of potential outcomes can be ordered conditional on an observed variable, which is called the monotone instrument. In my model, the monotone instrument corresponds to the indicator variable that specifies whether the proxy variable is greater than a cutoff. My approach differs from the monotone instrument variable approach in that the monotone instrument constructed from inside the model. This paper is also related to imperfect instrument literature where the instrument is not exogenous, but the researcher makes assumptions about the sign and magnitude of the correlation between the instrument and the error term. Under this additional information, one can construct bounds for the parameter of interest (Nevo and Rosen (2012)).

This paper also contributed to production function literature with multi-dimensional unobserved heterogeneity. Recently Demirer (2020) developed a method to identify production functions with two unobserved productivity terms, labor-augmenting

[1]The production function estimation literature goes back to Marschak and Andrews (1944), who first recognized the endogeneity problem. First attempts to address the endogeneity problem have used panel data methods (Mundlak and Hoch (1965), Mundlak (1961)). However, in practice, these methods do not give satisfactory answers, as summarized by Griliches and Mairesse (1995). See also Blundell and Bond (2000).

and Hicks-neutral production productivity. Another strand of literature uses a random coefficient model to introduce firm-level unobserved heterogeneity (Kasahara et al. (2015), Li and Sasaki (2017), Balat et al. (2019)). These papers assume a Cobb-Douglas production function with firm-specific elasticities and recover the distributions of the coefficients under different assumptions.

Notation. I use the notation $F_a(t)$ and $F_a(t \mid b)$ to donate the distribution of variable a and the distribution of a conditional on b, respectively. Similarly, I use $f_a(t)$ and $f_a(t \mid b)$ to denote the probability density function of random variable a and the probability density function of a conditional on b, respectively.

2 Model

In this section, I describe a production function model and then specify the assumptions. The model builds on the standard proxy variable framework introduced by OP, but allows for deviations from most of its strong assumptions. I discuss how my model differs from the proxy variable framework and the implications of the differences for identification. Even though the model is presented building on OP's framework, it can easily be adopted to other proxy variable estimation methods such as LP and ACF. I consider these extensions in Section 5.

2.1 Production Function

I consider a value-added Cobb-Douglas production function to demonstrate the main results of the paper. The identification strategy applies to other forms of production functions, which I will discuss in Section XX. The value-added Cobb-Douglas production function is given by

$$y_{it} = \theta_k k_{it} + \theta_l l_{it} + \omega_{it} + \epsilon_{it}, \tag{2.1}$$

where y_{it} denotes log-output, k_{it} denotes log-capital, and l_{it} denotes log-labor input. The model includes two unobserved productivity shocks, ω_{it} and ϵ_{it}. ω_{it} represents the *persistent* component of productivity, correlated over time; ϵ_{it} represents the *transitory* component of productivity that is independently and identically distributed. Firms observe ω_{it} before choosing inputs, whereas ϵ_{it} can be fully observed, partially observed, or not observed. These assumptions imply that both shocks can be correlated with inputs, which differs from the standard production function framework where ϵ_{it} is assumed to be not observed by the firm before making input decisions and is often considered as a measurement error in output.

The data consists of a panel of firms observed over periods $t = 1, \ldots, T$. Observations are independently and identically distributed across firms, but they can be serially correlated within the firm. The objective is to estimate the production function parameters, (θ_l, θ_k). I assume that capital is a dynamic input, meaning that the firm chooses capital one period in advance. As a result, capital is a state variable in the firm's dynamic optimization problem. The model is agnostic about the type of labor input. It can be static, dynamic, or partially dynamic, chosen between periods t and $t-1$ as considered in ACF. For the main text, I will assume that labor is a static input. Appendix XX demonstrates how to modify the identification strategy and estimation procedure when labor is a dynamic input.

2.2 Assumptions

My assumptions follow the structure of the proxy variable approach assumptions but relax them in several ways. This section presents the assumptions and describes the ways in which it is a less restrictive model than the proxy variable model. The first assumption defines the firm's information set.

Assumption 2.1 (Information Set)**.** *Let $\mathcal{I}_{it}$ denote firm i's information set at period t. I assume that past and current persistent productivity shocks are in firm's information set, that is, $\{\omega_{i\tau}\}_{\tau=-\infty}^{t} \in \mathcal{I}_{it}$. The transitory shocks satisfy $\mathbb{E}[\epsilon_{it} \mid \mathcal{I}_{it-1}] = 0$.*

This assumption distinguishes the roles of two productivity shocks. The persistent productivity, ω_{it}, is observed by the firm, so it is in the information set of the firm, but even if it is observed it is not informative about future productivity: $\mathbb{E}[\epsilon_{it} \mid \mathcal{I}_{it-1}] = 0$. Therefore, even if ϵ_{it} can be interpreted as a productivity shock it has limited dynamic implications.[2] The next assumption restricts the distribution of persistent productivity shock.

Assumption 2.2 (Markov Property)**.** *The persistent productivity shock follows an exogenous first-order Markov process.*

$$P(\omega_{it+1} \mid \mathcal{I}_{it}) = P(\omega_{it+1} \mid \omega_{it}),$$

and the distribution is stochastically increasing in ω_{it}.

[2]Note that since this includes mean independence, not full independence, the firm's dynamic decision can still be affected by ϵ_{it}, as the distribution of ϵ_{it} can give information about future production.

This assumption is standard in the literature and states that current productivity is the only information signal about future productivity (in the firm's information set). An implication of this assumption is that the transitory productivity shock ϵ_{it} is not informative about the distribution of future productivity shocks once we condition on ω_{it}. However, this assumption does not restrict the correlation between two productivity shocks, so they might be positively correlated. The second part of the assumption says that the distribution of future persistent productivity is stochastically increasing in ω_{it}. This assumption was first made by OP and indicates that more productive firms at the current period are likely to be more productive next period.[3] This assumption is critical for the moment inequality approach developed in this paper.

Assumption 2.3 (Materials Demand). *The firm's materials demand is given by,*

$$m_{it} = f_t(k_{it}, \omega_{it}, \xi_{it}),$$

where $\xi_{it} \in \mathbb{R}^L$ is a vector of unobserved random variables that affects firm's investment and it is assumed to be jointly independent of ω_{it} conditional on k_{it-1}.

According to this assumption, materials input depends on the two standard state variables, capital and persistent productivity, as well as other unobserved variables denoted by ξ_{it}. This assumption is the key deviation from the proxy variable approach, which assumes that the proxy variable depends only on capital and productivity shock. This assumption is made because with multi-dimensional unobserved heterogeneity, the key step in estimation, inverting the productivity shock, fails. Since my approach is different, I do not rely on invertibility.

Single-dimensional heterogeneity is a very restrictive assumption in firm models. It assumes that firms only consider their capital level and productivity when making their static input decisions. However, they are many variables that could affect a firm static input decision. Among them are input prices (both materials and other inputs), quality of inputs, and demand shocks. Since we typically do not observe these dimensions in the data, it is reasonable to model them as unobserved heterogeneity in the model. Also, note that we do not assume that ξ_{it} is single-dimensional, so ξ_{it} captures any of these shocks.

One important advantage of allowing for ξ is accommodating the rich structure of the product competition. Without ξ, the assumption is that there are no specific demand shocks in the model because if they are both own and competitor's demand

[3] A distribution is stochastically increasing if $P(\omega_{it+1} \mid \bar{\omega}_{it})$ first-order stochastically dominates $P(\omega_{it+1} \mid \tilde{\omega}_{it})$ if and only of $\bar{\omega}_{it} \geqslant \tilde{\omega}_{it}$.

shocks should affect the firm's materials demand. Therefore, implicitly, in prior approaches, identification is possible only under perfect competition or monopolistic competition with symmetric demand curves. This shows the importance of allowing multidimensional heterogeneity in a firm model to capture a richer competition structure. This is especially relevant when production function estimates are used for calculating markups (**de_loecker_rise_2018**).

Another condition in Assumption 2.3 is that ξ_{it} is independent of persistent productivity conditional on the last period's capital. Therefore, given the firm's capital level, the variables that affect its materials decision are not informative about the persistent productivity shock, ω_{it}. Although this assumption is restrictive, it allows for multi-dimensional heterogeneity and productivity. Some of the identification results presented later do not require this assumption, so it is still possible to make inferences on the parameters without this assumption. However, this assumption gives additional moment inequalities, which are likely to make the identified set tighter.

The assumption is particularly attractive if one wants to use investment rather than materials as a proxy variable. Investment is a dynamic decision that could be affected by many other variables. This could be heterogeneity in adjustment cost, heterogeneity in expectation, credit contract, and price for capital investments.[4] In case of investment proxy ξ_{it} will capture these potential sources of heterogeneity that could affect firms' investment decisions.

Assumption 2.3 also accommodates measurement error in capital, as one interpretation of ξ_{it} could be measurement error in investment. I discuss this point in Section 5.5, since measurement error in capital is an important concern in production function estimation.

Assumption 2.4 (Imperfect Proxy). *$f_t(k_{it-1}, \omega_{it}, \xi_{it})$ is weakly increasing in ω_{it} conditional on (k_{it-1}, ξ_{it}).*

This assumption relaxes the standard assumption of strictly increasing proxy variables in productivity to weak monotonicity. Single-dimensional unobserved heterogeneity and string monotonicity are the key assumptions in the proxy variable approach that allows one to invert the input demand function and essentially 'observe' the productivity using the proxy variable. Under my assumption, materials is no

[4]There is strong evidence for heterogeneity in adjustment cost. For example, Goolsbee and Gross (2000) present empirical evidence on heterogeneity in adjustment cost. Cooper and Haltiwanger (2006) argue that there is substantial heterogeneity in capital associated with heterogeneity in adjustment costs. Hamermesh and Pfann (1996), in a review paper, claim that heterogeneity in adjustment cost is a key source of heterogeneity across firms and should be included in models of firm behavior.

longer a 'perfect proxy' because the one-to-one relationship between investment and productivity no longer holds. However, by this assumption of weak monotonicity, materials demand is still informative about productivity because it could affect materials demand. For this reasons, in my framework materials become an imperfect proxy variable. My identification approach relies on capturing the information in an imperfect proxy variable via moment inequalities.

This assumption is particularly appealing when investment is used as a proxy variable. As observed by LP, strict monotonicity for investment is a particularly strong assumption because investment is often lumpy in the data. Moreover, firms do not invest in every period due to non-convex adjustment costs and option value of delaying investment (XX, XX) This suggests that investment is not continuous in productivity. OP drop firms with zero investment to overcome this problem. LP propose using materials instead of investment as a proxy. My approach is robust to both zero and lumpy investments in the data.

2.3 Discussion

To summarize, Assumptions 2.3 and 2.4 are the key differences of this paper from the standard assumptions, which assume that $m_{it} = f_t(k_{it}, \omega_{it})$ and $f_t(k_{it}, \omega_{it})$ is strongly increasing in ω_{it}. These assumptions limit the dimension of unobserved heterogeneity that impacts firm behavior. I relax these two strong assumptions on the functional form of investment by assuming that (i) investment is weakly increasing in productivity and (ii) there are other unobservables affecting the investment decision. Under these assumptions, f_t is not invertible, which is the key step in the proxy variable approach to control for unobserved productivity.

3 Identification

This section derives a set of conditional moment inequalities using the imperfect proxy based on the modeling assumptions. I derive moment inequalities and study identification under a wide range of assumptions from least restrictive to most restrictive.

3.1 Identification with Other Assumptions

The first step is deriving moment inequalities to establish stochastic ordering of productivity shocks in two set of firms based on the proxy variable. In particular, I will first study how productivity distributions of high 'high' and 'low' materials firms compare with each other. To do this, let me define $\tilde{\mathcal{X}} = \{(k, z) : 0 < \text{Prob}(i_{it} < z \mid k_{it} = k) < 1\}$. Here, $\tilde{\mathcal{X}}$ denotes the support of materials conditional on capital k_{it}. To simplify the exposition, I also assume that m_{it} has a continuous distribution function.

Proposition 3.1. *(MLRP) Assumptions 2.1-2.4, along with some regulatory conditions, imply that for $(k_{it-1}, z) \in \tilde{\mathcal{X}}$*

$$\frac{f_{\omega_{it}}(t \mid k_{it-1}, i_{it} > z)}{f_{\omega_{it}}(t \mid k_{it-1}, i_{it} < z)}$$

is increasing in t, that is, it satisfies the Monotone Likelihood Ratio Property (MLRP).

Proof. See Appendix A.

This proposition is the main result of that paper to establish the partial identification of production function parameters. It shows that a proxy variable that is weakly monotonic in productivity can provide a stochastic ordering of productivity distributions. Moreover, this proposition shows that MLRP property encompasses all the information given by a proxy variable that is weakly monotone in productivity.

Next, we provide two corollaries based on this result.

Corollary 3.1. *(MLRP) The distribution function of ω_{it} conditional on k_{it-1} and $\{i_{it} > z\}$ first order stochastically dominates (FOSD) the distribution function of ω_{it} conditional on k_{it-1} and $\{i_{it} < z\}$*

$$F_{\omega_{it}}(t \mid k_{it}, m_{it} > z) \geqslant F_{\omega_{it}}(t \mid k_{it}, m_{it} < z),$$

for $t \geqslant 0$ and $(k_{it}, z) \in \tilde{\mathcal{X}}$.

Corollary 3.2. *(Moment Inequality)The mean of ω_{it} conditional on k_{it-1} and $\{i_{it} > z\}$ is greater than the mean of ω_{it} conditional on k_{it-1} and $\{i_{it} < z\}$:*

$$\mathbb{E}[\omega_{it} \mid k_{it-1}, i_{it} > z] \geqslant \mathbb{E}[\omega_{it} \mid k_{it-1}, i_{it} < z], \tag{3.1}$$

for $(k_{it}, z) \in \tilde{\mathcal{X}}$.

Proof. Omitted A.

These results are direct implications of the MLRP property as MLRP implies first-order stochastic dominance. Later, we will see that operationalizing these different results on stochastic ordering depends on some other statistical assumptions. Therefore, we list these results separately and we will look at how we can use these under different assumptions to derive moment inequalities.

Remark 3.1 (Comparison to Proxy Variable Approach)**.** Single dimensional unobserved heterogeneity and strong monotonicity of investment in productivity allow OP to invert the investment function and recover productivity shock as $\omega_{it} = f_t^{-1}(i_{it}, k_{it})$. An invertible investment function means that one can control for unobserved productivity by conditioning on observables. My model relaxes the two necessary conditions for invertibility. First, I allow the investment function to be weakly monotone in productivity. Second, there are additional unobserved variables affecting investment. Therefore, we can no longer compare the productivity levels of two firms by comparing their investments under my assumptions. When a proxy is invertible, it is a perfect proxy because the firm's ranking in investment equals the ranking in productivity. This makes it possible to infer productivity using the proxy. In my model, materials is not a perfect proxy, so it is not possible to recover productivity from investment. However, by weak monotonicity, investment still provides information about productivity, so it becomes an imperfect proxy. My proposition shows that an imperfect proxy can be used to order productivity stochastically, rather than deterministically. In particular, Proposition 3.1 says that when firms are grouped based on how much they demand materials, we can infer that high materials firms will be more productive than low investment firms, on average. The main idea for identification is to use these stochastic orderings in the form of moment inequalities to set identify the production function.

Remark 3.2 (Conditioning on Two Investment Levels)**.** One might think that a moment inequality similar to Equation (3.1) holds, conditional on two different investment levels:

$$\mathbb{E}[\omega_{it} \mid k_{it}, m_{it} = z_1] \geqslant \mathbb{E}[\omega_{it} \mid k_{it}, i_{it} = z_2],$$

where $z_1 < z_2$. However, this inequality does not hold, as it is easy to find counterexamples under the modeling assumptions. Therefore, it is crucial to establish moment inequalities based on two sets of firms with high and low materials demand.

Remark 3.3 (Relation to Monotone Instrument and Imperfect Instrument Literature)**.** This paper is related to the literature on monotone instrumental variables (Manski (1997), Manski and Pepper (2000)). This literature assumes that the mean potential outcomes are ordered based on an observed variable, which is called a monotone instrument. In my model, investment can be considered a monotone instrument for productivity. The main difference of my model from the standard monotone instrumental variable approach is that the monotone instrument comes from within the model in this paper. My approach is also related to the 'imperfect instrument approach,' which assumes that the researcher has some prior information about the correlation between the endogenous variable and unobserved heterogeneity. This information is then used to construct moment inequalities. See, for example; Nevo and Rosen (2012) and Conley et al. (2012).

3.2 Derivation of Moment Inequalities

In this section, we will use the results in the previous section to derive moment inequalities. We will do this in two steps. First, we will derive inequalities for the labor elasticities, and in the second step, we will derive inequalities for both parameters. This second step procedure is similar to the estimation with proxy variable (LP), where the labor elasticity is identified first, and then capital elasticity is identified.

3.2.1 First Step

In the first step, we will estimate bounds for θ_l only. We define the following function which takes data and parameters as arguments:

$$m_1(w_{it}, \tilde{\theta}) := y_{it} - \tilde{\theta}_k k_{it} - \tilde{\theta}_l l_{it} = \omega_{it} + \epsilon_{it} \tag{3.2}$$

with $w_{it} = (k_{it}, l_{it})$ and $\tilde{\theta} = (\tilde{\theta}_l, \tilde{\theta}_k)$. Also, let θ denote the vector of true parameter values. By proposition (3.1) we have:

$$\mathbb{E}[\omega_{it} \mid k_{it}, m_{it} > z] \geqslant \mathbb{E}[\omega_{it} \mid k_{it}, m_{it} < z],$$

Substituting $m(w_{it}, \tilde{\theta})$ using Equation (3.4) we obtain

$$\mathbb{E}[m(w_{it}, \theta) \mid k_{it}, m_{it} > z] - \mathbb{E}[m(w_{it}, \theta) \mid k_{it}, m_{it} < z] \geqslant 0.$$

for $(k_{it}, z) \in \tilde{\mathcal{X}}$. Define $I^h_{it}(z) := \mathbb{1}\{m_{it} > z\}$ and $I^l_{it}(z) := \mathbb{1}\{m_{it} < z\}$. $I^h_{it}(z)$ equals 1 if a firm invests more than z at period t and zero otherwise. The opposite is true for $I^l_{it}(z)$. I call firms with $I^h_{it}(z) = 1$ as 'high' investment and $I^l_{it}(z) = 1$ as 'low' investment firms. With some abuse of notation, I treat these variables as events when they are in the conditioning set. Using this, we obtain

$$\theta_l \big(\mathbb{E}[l_{it} \mid I^h_{it}(z), k_{it}] - \mathbb{E}[l_{it} \mid I^l_{it}(z), k_{it}] \big) \leqslant \big(\mathbb{E}[y_{it} \mid I^h_{it}(z), k_{it}] - \mathbb{E}[y_{it} \mid I^l_{it}(z), k_{it}] \big).$$

Note that the capital elasticity parameter, θ_k, dropped from this moment inequality because we are conditioning on capital in the current period. This is analogue to the fact that in the proxy variable approach with static labor (as in LP) one identifies only labor coefficient in the first step. This gives us a bound for the labor elasticity in the following form:

$$\theta_l \leqslant \frac{\mathbb{E}[y_{it} \mid I^h_{it}(z), k_{it}] - \mathbb{E}[y_{it} \mid I^l_{it}(z), k_{it}]}{\mathbb{E}[l_{it} \mid I^h_{it}(z), k_{it}] - \mathbb{E}[l_{it} \mid I^l_{it}(z), k_{it}]}$$

or $(k_{it}, z) \in \tilde{\mathcal{X}}$. Therefore in this step, for each k_{it} and z we obtain bounds for the labor elasticity parameter. Whether this bound is an upper or lower bound depends on the sign of the nominator and denominator above.

3.2.2 Second Step

In the second step, we derive moment inequalities for capital and labor. To do this, we use the following representation of the production shock that follows from the Markov process:

$$\omega_{it} = g(\omega_{it-1}) + \zeta_{it},$$

where $\zeta_{it} = \omega_{it} - \mathbb{E}[\omega_{it} \mid \omega_{it-1}]$ and $\mathbb{E}[\zeta_{it} \mid I_{it-1}] = 0$ by construction. Also, the assumption that $P(\omega_{it} \mid \omega_{it-1})$ is stochastically increasing implies that $g(\omega_{it-1})$ is a monotone function. This representation of ω_{it} has been commonly used in the proxy variable approach for constructing moments. Substituting productivity into the production function yields:

$$y_{it} = \theta_k k_{it} + \theta_l l_{it} + g(\omega_{it-1}) + \zeta_{it} + \epsilon_{it}. \tag{3.3}$$

This representation of the production function involves three error terms: innovation to productivity ζ_{it}, measurement error in labor η^l_{it}, and the transitory productivity shock ϵ_{it}. Let me define a function that takes data and parameters:

$$m_2(w_{it}, \tilde{\theta}) := y_{it} - \tilde{\theta}_k k_{it} - \tilde{\theta}_l l_{it} \tag{3.4}$$

with $w_{it} = (k_{it}, l_{it})$ and $\tilde{\theta} = (\tilde{\theta}_l, \tilde{\theta}_k)$. Also let θ denote the vector of true parameter values. The next proposition presents a conditional moment inequality using Equation

(3.3) and Proposition 3.1. For $(k_{it-1}, z) \in \tilde{\mathcal{X}}$, we obtain

$$\mathbb{E}[m(w_{it}, \theta) \mid k_{it-1}, I_{it}^h(z)] - \mathbb{E}[m(w_{it}, \theta) \mid k_{it-1}, I_{it}^l(z)] \geqslant 0. \tag{3.5}$$

This proposition is the main identification result of the paper. Conditional on k_{it-1} if we compare two groups of firms, one with investment greater than z and one with investment lower than z, Equation (3.5) is satisfied at the true parameter values. The key conditions needed for this proposition are monotonicity of $g(\omega_{it-1})$ and the weak monotonicity of investment in productivity.

Our two-step approach is similar to the two steps estimation of the proxy variable approach. However, instead of identifying the labor elasticity in the first step, we constructed a bound for labor elasticity, and then the second step involves bounds for both labor and capital elasticity.

A necessary condition for this proposition to hold is that $(\zeta_{it}, \eta_{it}^l, \epsilon_{it})$ are orthogonal to the firm's information set at $t-1$. Recall that Proposition 3.1(iii) provides moment inequalities in terms of ω_{it}. However, we can only recover $g(w_{it-1}) + \zeta_{it} + \eta_{it}^l + \epsilon_{it}$ from the observed variables and parameters. Therefore, we need to account for $(\zeta_{it}, \eta_{it}^l, \epsilon_{it})$. The orthogonality condition allows me to achieve this, as $\zeta_{it} + \eta_{it}^l + \epsilon_{it}$ drop from the moment inequality in Equation (3.5) when we take conditional expectations.

3.3 Identified Set

In this section, I characterize the identified set using the derived moment inequalities. Since I have conditional moment inequalities, the identified set is given by intersection bounds. Recall that the true parameter satisfies:

$$\mathbb{E}[y_{it} - \theta_k k_{it} - \theta_l l_{it} \mid k_{it-1}, I_{it-1}^h(z)] - \mathbb{E}[y_{it} - \theta_k k_{it} - \theta_l l_{it} \mid k_{it-1}, I_{it-1}^l(z)] \geqslant 0 \tag{3.6}$$

$$\mathbb{E}[y_{it} - \theta_l l_{it} \mid k_{it}, I_{it}^h(z)] - \mathbb{E}[y_{it} - \theta_l l_{it} \mid k_{it}, I_{it}^l(z)] \geqslant 0 \tag{3.7}$$

I can write the moment inequality in Equation (3.6) as: This expression shows that the identified set depends on the mean capital and labor of 'high' and 'low' materials firms. For example, if average capital and labor do not vary with investment, then the identified set would not be informative. To make this more concrete, let me characterize the identified set. Define

$$a_l(z, k_t) := \mathbb{E}[l_{it} \mid k_t, I_{it}^h(z)] - \mathbb{E}[l_{it} \mid k_t, I_{it}^l(z)], \tag{3.8}$$

$$a_k(z, k_t) := \mathbb{E}[k_{it} \mid k_t, I_{it}^h(z)] - \mathbb{E}[k_{it} \mid k_t, I_{it}^l(z)], \tag{3.9}$$

$$a_y(z, k_t) := \mathbb{E}[y_{it} \mid k_t, I_{it}^h(z)] - \mathbb{E}[y_{it} \mid k_t, I_{it}^l(z)]. \tag{3.10}$$

Using these definitions, the moment inequality can be expressed as

$$\begin{aligned} a_k(z, k_{t-1})\theta_k + a_l(z, k_{t-1})\theta_l \leqslant y_l(z, k_{t-1}), \qquad (k, z) \in \tilde{\mathcal{X}} \\ a_l(z, k_t)\theta_l \leqslant y_l(z, k_t), \qquad (k, z) \in \tilde{\mathcal{X}} \end{aligned}$$

The identified set, conditional on k and z, is a region defined by a half-plane. Therefore, the identified set is the intersection of these half-planes.

Proposition 3.2 (Identified Set)**.** *Assume $\theta \in \tilde{\Theta}$, a compact parameter space. The identified set Θ is defined as the set of parameters that satisfy the conditional moment inequalities*

$$\Theta := \Bigg\{ \tilde{\theta} \in \tilde{\Theta} : \bigcap_{(k,z)\in\tilde{\mathcal{X}}} a_y(z, k_{t-1}) - \tilde{\theta}_k a_k(z, k_{t-1}) - \tilde{\theta}_l a_l(z, k_{t-1}) \geqslant 0, \\ a_y(z, k_t) - \tilde{\theta}_l a_l(z, k_{t-1}) \geqslant 0 \quad a.s. \Bigg\}, \tag{3.11}$$

and the identified set contains true parameter value $\theta \in \Theta$.

3.4 Moment Inequalities Using FOSD and MLRP

Proposition 3.1 establishes that the distributions of productivity conditional on high and low investment satisfy MLRP and FOSD. However, when characterizing the identified set, I only used the mean ordering, a weaker implication of MLRP and FOSD. This is because even though MLRP and FOSD hold for ω_{it} conditional on high and low investment, I can only recover $g(w_{it-1}) + \zeta_{it} + \eta_{it}^l + \epsilon_{it}$ using the data and parameters. Therefore, I need additional conditions that ensure that MLRP and FOSD are preserved when there are additive errors. The following two theorems present those conditions.

Theorem 3.1 (Shaked and Shanthikumar (2007))**.** *Let X_1 and X_2 be two independent random variables, and Y_1 and Y_2 be another two independent random variables. If*

$X_i \leqslant_{FOSD} Y_i$ for $i = 1, 2$ then

$$X_1 + X_2 \leqslant_{FOSD} Y_1 + Y_2.$$

Theorem 3.2 (Shaked and Shanthikumar (2007))**.** *Let X_1, X_2 and Z be random variables such that X_1 and Z are independent and X_2 and Z are independent. If $X_1 \leqslant_{MLRP} X_2$ and Z has a log-concave probability density functions, then*

$$X_1 + Z \leqslant_{MLRP} X_2 + Z.$$

These two theorems suggest that in order for MLRP and FOSD to be preserved under convolutions I need: (i) an independence condition for FOSD, and (ii) independence and log-concavity conditions for MLRP. Therefore, I next impose these conditions on the unobservables to derive moment inequalities using FOSD and MLRP.

3.4.1 Identified Set Using FOSD

As Theorem 3.1 suggests, I need to impose independence restrictions on unobservables to preserve FOSD ordering.

Assumption 3.1. *$(\eta_{it}^l, \zeta_{it}, \epsilon_{it})$ are jointly independent from $\mathcal{I}_{it-1}$.*

With this assumption ω_{it-1} becomes jointly independent from the rest of the unobservables, $\zeta_{it} + \eta_{it}^l + \epsilon_{it}$, conditional on $\mathcal{I}_{it-1}$. Therefore, MLRP for $g(\omega_{it-1})$ conditional on high and low investment is preserved for $g(\omega_{it-1}) + \zeta_{it} + \epsilon_{it} + \eta_{it}$ conditional on high and low materials demand. Under this assumption, I can strengthen the mean ordering in Equation (3.1) to the first-order stochastic dominance ordering, and characterize the identified set accordingly.

Proposition 3.3 (Identified Set-FOSD)**.** *Under Assumptions 2.1-2.4 and Assumption 3.1, the true parameter $\theta \in \tilde{\Theta}$ satisfies the following condition*

$$\begin{aligned} F_{m_1(w_{it},\theta)}(\cdot \mid k_{it}, I_{it}^h(z) = 1) - F_{m_{[1]}(w_{it},\theta)}(\cdot \mid k_{it}, I_{it}^l(z) = 1) \geqslant 0, \\ F_{m_2(w_{it},\theta)}(\cdot \mid k_{it}, I_{it}^h(z) = 1) - F_{m_{[2]}(w_{it},\theta)}(\cdot \mid k_{it}, I_{it}^l(z) = 1) \geqslant 0 \end{aligned} \tag{3.12}$$

for all $(k_{it}, z) \in \tilde{\mathcal{X}}$. The identified set (conditional on k and z) is

$$\Theta_t^{FOSD} = \{\widehat{\theta} \in \Theta : (3.12) \text{ holds with } \theta \text{ in place of } \widehat{\theta}\}.$$

The independence assumption is non-standard in production function models, but it is difficult to imagine situations where mean independence holds and independence does not hold. Next, I specify the assumption that allows me to use the MLRP result in Proposition 3.1(i) for identification.

3.4.2 Identified Set Using MLRP

As Theorem 3.2 suggests, I need to impose independence and shape restrictions on the distributions of unobservables to be able to use MLRP.

Assumption 3.2. *$(\eta_{it}^l, \zeta_{it}, \epsilon_{it})$ are jointly independent from $\mathcal{I}_{it-1}$, and each variable in $\eta_{it}^l, \zeta_{it}, \epsilon_{it}$ has a log-concave probability distribution function.*

Under this assumption, I can strengthen the mean inequality in Proposition (5.2) to MLRP and characterize the identified set accordingly.

Proposition 3.4 (Identified Set-MLRP)**.** *Under Assumptions 2.1-**??** and Assumption 3.2, the true parameter $\theta \in \tilde{\Theta}$ satisfies the following inequality:*

$$\begin{aligned}
&F_{m_1(w_{it},\theta)}(\cdot \mid a \leqslant m_1(\cdot) \leqslant b, k_{it} = k, I_{it}^h(z)) - F_{m_1(w_{it},\theta)}(\cdot \mid a \leqslant m_1(\cdot) \leqslant b, k_{it}, I_{it}^l(z)) \geqslant 0 \\
&F_{m_2(w_{it},\theta)}(\cdot \mid a \leqslant m_1(\cdot) \leqslant b, k_{it} = k, I_{it}^h(z)) - F_{m_{2(w_{it},\theta)}}(\cdot \mid a \leqslant m_1(\cdot) \leqslant b, k_{it}, I_{it}^l(z)) \geqslant 0
\end{aligned} \tag{3.13}$$

for all $(k_{it}, z) \in \tilde{\mathcal{X}}$ and for all (a, b) such that $a < b$. The identified set (conditional on (k, z)) is

$$\Theta_t^{MLRP} = \{\widehat{\theta} \in \Theta : (3.13) \textit{ holds with } \widehat{\theta} \textit{ in place of } \theta\}.$$

Most well known distributions, such as those in the exponential family, satisfy the assumptions required for this proposition.

3.5 Discussion

An advantage of identification analysis under a wide range of assumptions is that we can see the role of each assumption in identification. For example, we can start with the most general model to impose as few restrictions as possible. If the estimated set is not informative, then a nested model can be considered to shrink the identified set. Also, comparing estimates from a nested and a nesting model would test the restrictions imposed by the nested model.

Note also that proxy variable specification is a special case of my framework. So, if the estimates set is not information, one can use the proxy variable approach to point identify the parameters. It is also worth noting that, the identified set under my assumptions does not have to include the point estimates obtained from proxy variable method. The reason is that under the proxy variable assumptions the model is over-identified. If my partial identification method uses overidentification restrictions, the point estimates might be excluded from the identified set. This would mean rejecting the proxy variable assumptions.

One may think that set identifying the production function parameters is not useful unless the set is tight. As in most set identification results, the informativeness of the identified set depends on the data and empirical setting. However, as discussed above, there are other advantages of using my framework. Most importantly, since the standard OP approach is nested under my assumptions, there is no harm in starting with a more general model.

4 Estimation

The identification results generate conditional moment inequalities conditions on high and low materials demand. The estimation procedure requires conducting inference using these moment inequalities. To achieve this, I will define the propensity of high and low materials conditional on k_{it} and integrate that into moment equalities.

First, I define the propensity score, which equals the probability that materials are greater than a cutoff z as

$$m(k_{it}, z) = \mathbb{E}[I_{it}^h(z) \mid k_{it}].$$

I define the moments using the propensity scores in the following ways.

$$\begin{aligned}
\mathbb{E}[\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) \mid k_{it}, I_{it}^h(z)] &= \mathbb{E}\left[\frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) I_{it}^h(z)}{m(k_{it}, z)} \mid k_{it}\right], \\
\mathbb{E}[\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) \mid k_{it}, I_{it}^l(z)] &= \mathbb{E}\left[\frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big)\big(1 - I_{it}^h(z)\big)}{\big(1 - m(k_{it}, z)\big)} \mid k_{it}\right].
\end{aligned}$$

Integrating out k_i, the unconditional moment inequality can be written as:

$$\mathbb{E}\left[\frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) I_{it}^h(z)}{m(k_{it}, z)} - \frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big)\big(1 - I_{it}^h(z)\big)}{1 - m(k_{it}, z)} \mid k_{it}\right] \geqslant 0.$$

Define

$$s^h(\theta, z) = \mathbb{E}\left[\frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) I_{it}^h(z)}{m(k_{it}, z)}\right], \qquad s^l(\theta, z) = \mathbb{E}\left[\frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big)\big(1 - I_{it}^h(z)\big)}{1 - m(k_{it}, z)}\right].$$

Estimation can be carried out by testing the hypothesis $s^h(\theta, z) \geqslant s^l(\theta, z)$ and inverting the test. Specifically, for a given θ, test $s^h(\theta, z) \geqslant \theta^l(\theta, z)$ for all z and include θ in the identified set if the the null hypothesis fails to be rejected. The natural estimators for $s^h(\theta, z)$ and $s^l(\theta, z)$ are

$$\widehat{s}^h(\theta, z) = \sum_{i=1}^{N} \frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) I_{it}^h(z)}{\widehat{m}(k_{it}, z)}, \qquad s^l(\theta, z) = \sum_{i=1}^{N} \frac{\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) I_{it}^l(z)}{1 - \widehat{m}(k_{it}, z)}, \tag{4.1}$$

where $\widehat{m}(k_{it}, z)$ is an estimate of $m(k_{it}, z)$. These functions involve the propensity score function, which is a nuisance function and needs to be estimated in the first stage. To make the estimation procedure more robust to an estimation error in the nuisance function, I can define the doubly robust moment functions. This requires other nuisance functions in the moment functions. Define

$$g^h(k, z, \theta) = \mathbb{E}[\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) \mid k_{it} = k, I_{it}^h(z)],$$
$$g^l(k, z, \theta) = \mathbb{E}[\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) \mid k_{it} = k, I_{it}^l(z)].$$

Using these functions, the doubly robust moments are

$$s_{db}^h(\theta, z) = g^h(k_{it}, z, \theta) + \frac{I_{it}^h(z)\Big(\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) - g^h(k_{it}, z, \theta)\Big)}{m(k_{it}, z)},$$
$$s_{db}^l(\theta, z) = g^l(k_{it}, z, \theta) + \frac{\big(1 - I_{it}^h(z)\big)\big((y_{it} - \theta_k k_{it} - \theta_l l_{it}) - g^l(k_{it}, z, \theta)\big)}{1 - m(k_{it}, z)}.$$

Using these doubly robust moments, we have the following conditional moment inequality:

$$\mathbb{E}[s_{db}^h(\theta, z) - s_{db}^l(\theta, z) \mid k_{it}] \geqslant 0.$$

The sample analog of these moments can be obtained similarly to Equation (4.1). Expectations of the doubly robust moments and the original moments are equal to each other at the true nuisance functions. Doubly robust moments have the property that if one of the nuisance functions is correct, then the moment is correct.

Finally, this moment condition implies that for a nonnegative w function, by the Law of Iterated Expectations

$$\mathbb{E}\Big[\big(s^h_{db}(\theta, z) - s^l_{db}(\theta, z)\big)w(k_{it})\Big] \geqslant 0.$$

Semenova (2017) studies moment inequality estimation with nuisance functions shows how to do inference when the nuisance functions are estimated using machine learning methods. One can also consider using conditional moment inequalities to tighten the identified set instead of integrating out capital. For that estimation problem one can use many moment inequalities framework of Chernozhukov et al. (2018) or conditional moment inequality estimation framework of Andrews and Shi (2013).

5 Extensions

The approach developed in this paper can be extended to other forms of production functions. To give some examples, I discuss the application to gross production function and using materials as the proxy variable instead of investment. I also show how my model accommodates measurement error in capital.

5.1 Gross Production Function

In this subsection, I show how to extend my model to a gross production function. The estimation procedure remains the same, with an increase in the number of parameters. A gross production function is given by

$$y_{it} = \theta_k k_{it} + \theta_l l_{it} + \theta_m m_{it} + \omega_{it} + \eta^l_{it} + \epsilon_{it}. \tag{5.1}$$

Similar to the main model, using the Markov assumption we can express the production function as

$$y_{it} = \theta_k k_{it} + \theta_l l_{it} + \theta_m m_{it} + g(\omega_{it-1}) + \zeta_{it} + \eta^l_{it} + \epsilon_{it}.$$

All the propositions presented for the main model apply to this model because they do not depend on the functional form of the production function. Therefore, we can construct a moment function as

$$m(w_{it}, \tilde{\theta}) := y_{it} - \tilde{\theta}_k k_{it} - \tilde{\theta}_l l_{it} - \tilde{\theta}_m m_{it},$$

with $w_{it} = (k_{it}, l_{it}, m_{it})$, $\tilde{\theta} = (\tilde{\theta}_l, \tilde{\theta}_k, \tilde{\theta}_m)$. Let θ denote the vector of true parameter values.

Proposition 5.1. *For all* $(k_{it}, z) \in \tilde{\mathcal{X}}$

$$\mathbb{E}[m(w_{it}, \theta) \mid k_{it}, i_{it-1} > z] - \mathbb{E}[m(w_{it}, \theta) \mid k_{it}, i_{it-1} < z] \geqslant 0.$$

Proof. Omitted.

This moment inequality can be used to estimate the model parameters. The only difference is that there are more parameters to estimate, so the estimated bounds might be less informative.

5.2 Using Materials as an Imperfect Proxy

Levinsohn and Petrin (2003) propose using materials instead of investment as a proxy for productivity shock motivated by the fact that investment is often lumpy in the data. Even though my framework is robust to this problem, this paper allows for using materials as a proxy instead of investment. To show this, I need to replace Assumption 2.3 with the following assumption.

Assumption 5.1. *Firms' materials decision is given by*

$$m_{it} = m_t(k_{it-1}, \omega_{it}, \xi_{it}),$$

where ξ_{it} *is a vector of unobserved random variables that affect the firm's materials decision and it is assumed to be independent of* ω_{it} *conditional on* k_{it-1}*.*

When materials is used as an imperfect proxy, the moment inequalities take the form in the following proposition.

Proposition 5.2. *For all* $(k_{it}, z) \in \tilde{\mathcal{X}}$*, the true parameter value,* θ*, satisfies*

$$\mathbb{E}[m(w_{it}, \theta) \mid k_{it-1}, m_{it} > z] - \mathbb{E}[m(w_{it}, \theta) \mid k_{it-1}, m_{it} < z] \geqslant 0.$$

The moment function, $m(w_{it}, \theta)$, is defined in Equation (3.4). This moment inequality can be used to estimate the model parameters.

5.3 Identification with Predetermined Capital

To accommodate predetermined capital, the standard timing assumption in proxy variable framework, I replace Assumption **??** with the following assumption.

Assumption 5.2. *Capital accumulates according to*

$$k_{it} = \delta k_{it-1} + i_{it-1}.$$

This assumption implies that the amount of capital used for time t production is determined at time $t-1$. I also need to replace Assumption (2.3) with the following assumption.

Assumption 5.3. *Firms' investment decision is given by*

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

where ξ_{it} is a vector of unobserved variables that affect firm's investment decision and it is assumed to be independent of ω_{it} conditional on k_{it}.

These changes in the assumptions affect only the conditioning set in moment inequalities. In particular, I need to condition on (k_{it-1}, i_{it-1}) instead of (k_{it}, i_{it-1}). So the moment inequality becomes:

$$\mathbb{E}[m(w_{it}, \theta) \mid k_{it-1}, i_{it-1} > z] - \mathbb{E}[m(w_{it}, \theta) \mid k_{it-1}, i_{it-1} < z] \geqslant 0,$$

where the moment function, $m(w_{it}, \theta)$, is defined in Equation (3.4). This moment inequality can be used to estimate the parameters. This modification does not affect the remained of the estimation.

5.4 Deviations from Cobb-Douglas Functional Form

In my identification strategy, the Cobb-Douglas functional form is necessary only to account for measurement errors. Thus, if I rule out measurement errors, I can use a nonlinear production function known up to a finite dimensional parameter vector. To demonstrate this extension, consider the following model:

$$y_{it} = r(\theta, w_{it}) + \omega_{it} + \epsilon_{it}, \tag{5.2}$$

where $r(\theta, w_{it})$ is a known function but the parameter vector, θ, is unknown. For this model, the moment function becomes:

$$m^*(w_{it}, \tilde{\theta}) := y_{it} - r(\tilde{\theta}, w_{it}). \tag{5.3}$$

With this moment function, the results for the main model can be applied to estimate Equation (5.2).

5.5 Measurement Error in Capital

My framework can also accommodate measurement error in capital, which is important because among all inputs, capital is most prone to measurement error.

Let η_{it}^k denote measurement error in investment. It is natural to model measurement error in capital using measurement error in investment because capital is accumulated through investment. Also, capital is often constructed from investment series in the data.[5] Let i_{it}^* denote true investment. The observed investment is given by:

$$\begin{aligned} i_{it} &= i_{it}^* + \eta_{it}^k \\ &= f_t(k_{it-1}, \omega_{it}, \xi_{it}) + \eta_{it}^k. \end{aligned}$$

Measurement error, η_{it}^k, can be included into f_t function as a part of ξ_{it} vector. Define $\xi_{it}^* = (\xi_{it}, \zeta_{it}^k)$ and rewrite the investment function as

$$i_{it} = f_t(k_{it-1}, \omega_{it}, \xi_{it}^*).$$

With measurement error in investment, observed capital takes the form

$$k_{it} = k_{it}^* + \sum_{j=0}^{t}(1-\delta)^j \eta_{it}^k,$$

where k_{it}^* is the true capital observed by the firm. Substituting this into the production function yields

$$y_{it} = \theta_k k_{it} + \theta_l l_{it} + g(\omega_{it-1}) + \zeta_{it} - \left(\sum_{j=0}^{t}(1-\delta)^j \eta_{it}^k\right)\theta_k + \eta_{it}^l + \epsilon_{it}. \tag{5.4}$$

I can combine the measurement errors in capital and labor as $\eta_{it} = -\left(\sum_{j=0}^{t}(1-\delta)^j \eta_{it}^k\right)\theta_k + \eta_{it}^l$. With the combined measurement error, production function becomes:

$$y_{it} = \theta_k k_{it} + \theta_l l_{it} + g(\omega_{it-1}) + \zeta_{it} + \eta_{it} + \epsilon_{it}. \tag{5.5}$$

This equation takes the form I analyzed for my partial identification results. So if we

[5]For example, the perpetual inventory method.

assume that η_{it}^k be is independent of ω_{it} conditional on k_{it-1} and $\mathbb{E}[\eta_{it} \mid \mathcal{I}_{it-1}] = 0$, we can use the same moment inequalities derived in Equation (3.5).

Since measurement error in capital is an important problem, the literature has paid particular attention to measurement error in capital. Some examples are Hu et al. (2011), Collard-Wexler and De Loecker (2016), and Kim et al. (2016). These papers require an instrumental variable or another proxy variable to control for measurement error in capital and show point identification. In contrast, my method does not require another variable, but it gives a bound instead of a point.

6 Empirical Application

In this section, I apply my method to a panel production data from Turkey.

6.1 Data

The data for Turkey are provided by the Turkish Statistical Institute (TurkStat; formerly known as the State Institute of Statistics, SIS), which collects plant-level data for the manufacturing sector.[6] Periodically, Turkstat conducts the Census of Industry and Business Establishments (CIBE), which collects information on all manufacturing plants in Turkey. In addition, TurkStat conducts the Annual Surveys of Manufacturing Industries (ASMI) that covers all establishments with at least 10 employees. The set of establishments used for ASMI is obtained from the CIBE. In non-census years, the new private plants with at least 10 employees are obtained from the chambers of industry.

I use a sample from Annual Surveys of Manufacturing Industries, covering a period from 1983 to 2000. Data from a more recent period is available, but due to major changes in the survey methodology, it is not possible to link this dataset to the data from a more recent period. I limit the sample to only private establishments. I focus on the textile industry, which is the largest 3-digit industry in terms of the number of firms. My sample includes 1437 firms and 14271 year-firm observations.

The data includes gross revenue, investment, the book value of capital, materials expenditures and the number of production and administrative workers. The real value of annual output is obtained by deflating the plant's total annual sales revenues by an industry-specific price index. Material inputs include all purchases of intermediate inputs. I deflate the nominal value of total material input cost by each plant

[6] This dataset has been used by Levinsohn (1993) and Taymaz and Yilmaz (2015).

using the industry-level intermediate input price index. Finally, capital stock series is constructed using the perpetual inventory method where investment in new capital is combined with deflated capital from period $t-1$ to form capital in period t.

6.2 Empirical Model

For my empirical application, I consider a value-added production function.

$$y_{it} = \theta_k k_{it} + \theta_l l_{it} + \omega_{it} + \epsilon_{it}.$$

I maintain the standard timing assumption that capital is chosen one period in advance.[7] I use the following moment inequality for estimation

$$\mathbb{E}[y_{it} - \tilde{\theta}_k k_{it} - \tilde{\theta}_l l_{it} \mid k_{it-1}, i_{it-1} > z] - \mathbb{E}[y_{it} - \tilde{\theta}_k k_{it} - \tilde{\theta}_l l_{it} \mid k_{it-1}, i_{it-1} < z] \geqslant 0. \quad (6.1)$$

I follow the steps outlined in the estimation section to construct unconditional moment inequalities from this conditional moment inequality. In particular, I first calculate the empirical analog of $\widehat{s}^h_{db}(\theta, z)$ and $s^l_{db}(\theta, z)$ using

$$\widehat{s}^h_{db}(\theta, z) = \sum_{i=1}^{N} \widehat{g}^h_t(k_{it-1}, z, \theta) + \frac{I^h_{it}(z)\Big(\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) - \widehat{g}^h_t(k_{it-1}, z)\Big)}{1 - \widehat{m}_t(k_{it-1}, z)}, \quad (6.2)$$

$$\widehat{s}^l_{db}(\theta, z, \theta) = \sum_{i=1}^{N} \widehat{g}^l_t(k_{it-1}, z) + \frac{I^l_{it}(z)\Big(\big(y_{it} - \theta_k k_{it} - \theta_l l_{it}\big) - \widehat{g}^l_t(k_{it-1}, z)\Big)}{\widehat{m}_t(k_{it-1}, z)}, \quad (6.3)$$

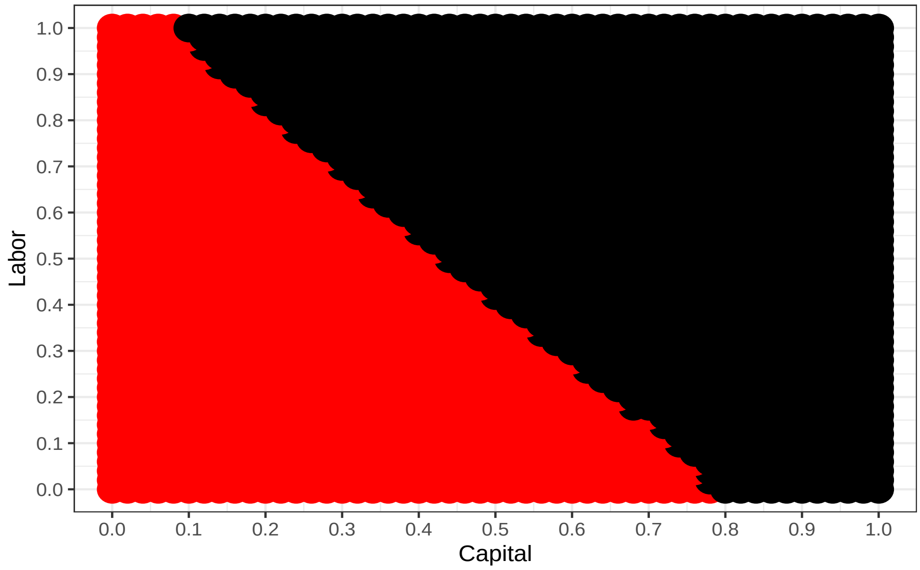
where I estimate the nuisance functions $\widehat{m}_t$, $\widehat{g}^h_t$ and $\widehat{g}^l_t$ using random forest method in the first stage. Following Semenova (2017), I also employ cross-fitting, i.e, I estimate the nuisance functions in the first half of the sample and construct the moments, $\widehat{s}^h_{db}(\theta, z)$ and $\widehat{s}^l_{db}(\theta, z)$, using the second half of the sample. I then swap these samples to avoid loss of efficiency.

I test the moment inequality $s^h_{db}(\theta, z) \geqslant s^l_{db}(\theta, z)$ using the Chernozhukov et al. (2018) many moment inequalities framework with the empirical analogs given in Equations (6.2) and (6.3). I obtain 50 moment inequalities by choosing 10 different z values from the support and testing moment inequalities in 5 different periods (84-87, 88-90, 91-93, 94-96, 97-00) for each z value.[8] The estimated set is constructed by inverting this test, that is, estimated set includes all parameter values for which I fail to reject

[7] I assume this because the capital series is constructed under this assumption using the perpetual inventory method.

[8] To calculate the critical values I use their multiplier bootstrap method.

Figure 1: Estimated Set



Note: The estimated of set for the value-added Cobb-Douglas production function parameters. The reported set shown in black covers the true parameter with 95% probability.

$s^h_{db}(\theta, z) \geqslant s^l_{db}(\theta, z)$.

I present the estimation results in Figure 1, where the red region shows the estimated set (with 95% coverage). As suggested in the identification analysis of Section 3, the estimated set is a half-plane. It is more informative about θ_k than θ_l, as it excludes the large values of capital elasticity values. One can also combine this estimated set with a priori restrictions on the production function based on economic theory. For example, if we impose that the restriction that the production function is constant returns to scale, then we can conclude from the identified set that the elasticity of capital is less than 0.28.

Even though the identified set does not give a definite answer about the production function parameters, it suggests that less restrictive assumptions in the proxy variable framework is still informative about production technology. Also, the estimated set is not sharp, since the estimation procedure does not use moment inequalities conditional on capital. Moreover, one can also use the moment inequalities derived from FOSD and MLRP restrictions under slightly stronger assumptions in Subsection 3.4. These estimations might give more informative estimated sets.

7 Conclusion

This paper extends the production function estimation literature by relaxing the restrictive assumptions of the proxy variable approach and showing that the parameters remain partially identified. My model (i) allows for multi-dimensional unobserved heterogeneity, (ii) relaxes strict monotonicity to weak monotonicity, (iii) accommodates a more general timing assumption. Also, the method is robust to measurement errors in inputs, an important problem in production function estimation.

I accomplish this by using an 'imperfect proxy' variable for identification. An 'imperfect proxy' variable contains information about productivity, but it cannot directly be used to control for productivity as in the proxy variable approach. Instead, an imperfect proxy variables generates stochastic orderings of productivity distributions, which can be exploited for estimation. I show how to use this stochastic ordering in the form of moment inequalities to obtain bounds for production function parameters.

A Proofs

A.1 Proof of Proposition B.1

Under Assumption 2.2 productivity shock can be written as

$$\omega_{it} = g(\omega_{it-1}) + \xi_{it},$$

where $g(\cdot)$ is a monotone function because the distribution is stochastically increasing in ω_{it} by Assumption 2.2. Substituting this into Equation (B.4) and expanding the left-hand side yield

$$\begin{aligned}
&\mathbb{E}\Big[\big(\omega_{it} + \epsilon_{it} - \omega_{it-1} - \epsilon_{it-1}\big)^2\Big] \\
&\qquad = \mathbb{E}\Big[\big(g(\omega_{it-1}) + \xi_{it} + \epsilon_{it} - \omega_{it-1} - \epsilon_{it-1}\big)^2\Big], \\
&\qquad = \mathbb{E}\Big[\big((g(\omega_{it-1}) - \omega_{it-1}) + \xi_{it} + \epsilon_{it} - \epsilon_{it-1}\big)^2\Big], \\
&\qquad = \mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{it-1}\big)^2\Big] + \mathbb{E}\Big[(\xi_{it} + \epsilon_{it} - \epsilon_{it-1}\big)^2\Big] + 2\mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{it-1}\big)(\xi_{it} + \epsilon_{it} - \epsilon_{it-1})
\end{aligned}$$

Expanding the right-hand side similarly

$$\begin{aligned}
&\mathbb{E}\Big[\big(\omega_{it} + \epsilon_{it} - \omega_{jt-1} - \epsilon_{jt-1}\big)^2\Big] \\
&= \mathbb{E}\Big[\big(g(\omega_{it-1}) + \xi_{it} + \epsilon_{it} - \omega_{jt-1} - \epsilon_{jt-1}\big)^2\Big] \\
&= \mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{jt-1} + \xi_{it} + \epsilon_{it} - \epsilon_{jt-1}\big)^2\Big] \\
&= \mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{jt-1}\big)^2\Big] + \mathbb{E}\Big[(\xi_{it} + \epsilon_{it} - \epsilon_{jt-1}\big)^2\Big] + 2\mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{jt-1}\big)(\xi_{it} + \epsilon_{it} - \epsilon_{jt-1})\Big].
\end{aligned}$$

Since observations are independently and identically distributed across firms and $\mathbb{E}[\epsilon_{it} \mid \mathcal{I}_{it-1}] = 0$, the second expectations are equal to each other

$$\mathbb{E}\Big[(\xi_{it} + \epsilon_{it} - \epsilon_{it-1}\big)^2\Big] = \mathbb{E}\Big[(\xi_{it} + \epsilon_{it} - \epsilon_{jt-1}\big)^2\Big]$$

Since $(\epsilon_{it}, \epsilon_{it-1}, \xi_{it})$ is orthogonal to ω_{it-1} and observations and independent and identically distributed, the third expectations equal zero:

$$\mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{jt-1}\big)(\xi_{it} + \epsilon_{it} - \epsilon_{jt-1})\Big] = \mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{it-1}\big)(\xi_{it} + \epsilon_{it} - \epsilon_{it-1})\Big] = 0$$

Therefore, to prove the proposition, I need to show that

$$\mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{it-1}\big)^2\Big] \leqslant \mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{jt-1}\big)^2\Big].$$

Expanding both sides

$$\begin{aligned}\mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{it-1}\big)^2\Big] &= \mathbb{E}[g(\omega_{it-1})^2] + \mathbb{E}[\omega_{it-1}^2] - 2\mathbb{E}[g(\omega_{it-1})\omega_{it-1}] \\ \mathbb{E}\Big[\big(g(\omega_{it-1}) - \omega_{jt-1}\big)^2\Big] &= \mathbb{E}[g(\omega_{it-1})^2] + \mathbb{E}[\omega_{jt-1}^2] - 2\mathbb{E}[g(\omega_{it-1})\omega_{jt-1}]\end{aligned}$$

The second and third moments are equal to each other by iid assumption

$$\mathbb{E}[g(\omega_{it-1})^2] = \mathbb{E}[g(\omega_{it-1})^2], \qquad \mathbb{E}[\omega_{it-1}^2] = \mathbb{E}[\omega_{jt-1}^2].$$

So I need show that

$$\mathbb{E}[g(\omega_{it-1})\omega_{it-1}] \geqslant \mathbb{E}[g(\omega_{it-1})\omega_{jt-1}].$$

Observe that $\mathbb{E}[g(\omega_{it-1})\omega_{jt-1}] = 0$. We also have $\mathbb{E}[g(\omega_{it-1})\omega_{it-1}] \geqslant 0$, because for a random variable X, $\mathbb{E}[f(X)X] \geqslant 0$ for an increasing function f. This gives the inequality in the proposition.

$$\mathbb{E}\Big[\big(\omega_{it} + \epsilon_{it} - \omega_{it-1} - \epsilon_{it-1}\big)^2\Big] \leqslant \mathbb{E}\Big[\big(\omega_{it} + \epsilon_{it} - \omega_{jt-1} - \epsilon_{jt-1}\big)^2\Big].$$

A.2 Proof of Proposition 3.1

Using the Bayes rule for continuous random variables, I write the conditional probability distribution function of ω_{it} as (by changing the notation slightly)

$$f(\omega_{it} \mid k_{it-1}, i_{it} > z) = \frac{\Pr(i_{it} > z \mid k_{it-1}, \omega_{it}) f(\omega_{it} \mid k_{it-1}) f(k_{it-1})}{\Pr(i_{it} > z \mid k_{it-1}) f(k_{it-1})} = \frac{\Pr(i_{it} > z \mid k_{it-1}, \omega_{it}) f(\omega_{it} \mid k_{it-1})}{\Pr(i_{it} > z \mid k_{it-1})}$$

and similarly for $f(\omega_{it} \mid k_{it}, i_{it} < z)$. By taking the ratio of the two

$$\begin{aligned}\frac{f(\omega_{it} \mid k_{it-1}, i_{it} > z)}{f(\omega_{it} \mid k_{it-1}, i_{it} < z)} &= \frac{\Pr(i_{it} > z \mid k_{it-1}, \omega_{it}) \Pr(i_{it} < z \mid k_{it-1})}{\Pr(i_{it} < z \mid k_{it-1}, \omega_{it}) \Pr(i_{it} > z \mid k_{it-1})} \\ &= \frac{\Pr(i_{it} > z \mid k_{it-1}, \omega_{it}) \Pr(i_{it} < z \mid k_{it-1})}{\big(1 - \Pr(i_{it} > z \mid k_{it-1}, \omega_{it})\big)\big(1 - \Pr(i_{it} < z \mid k_{it-1})\big)}\end{aligned}$$

This function is increasing in $\Pr(i_{it} > z \mid k_{it-1}, \omega_{it})$ because the numerator is increasing and denominator is decreasing in $\Pr(i_{it} > z \mid k_{it-1}, \omega_{it})$. So if I can show that $\Pr(i_{it} >$

$z \mid k_{it-1}, \omega_{it})$ is weakly increasing in ω_{it} that imply that

$$\frac{f(\omega_{it} \mid k_{it-1}, i_{it} > z)}{f(\omega_{it} \mid k_{it-1}, i_{it} < z)}$$

is a weakly increasing function of ω_{it}, and there the monotone likelihood ratio property holds. The next step is to show that $\Pr(i_{it} > z \mid k_{it}, \omega_{it})$ is weakly increasing in ω_{it}. I write $\Pr(i_{it} > z \mid k_{it}, \omega_{it})$ as

$$\begin{aligned}\Pr(i_{it} > z \mid k_{it-1}, \omega_{it}) &= \int \mathbb{1}\{f(k_{it-1}, \omega_{it}, \xi_{it}) > z\} f(\xi_{it} | \omega_{it}, k_{it-1}) d\xi_{it}, \\ &= \int \mathbb{1}\{f(k_{it-1}, \omega_{it}, \xi_{it}) > z\} f(\xi_{it} | k_{it-1}) d\xi_{it},\end{aligned}$$

where the second line follows from Assumption 2.3. Since by assumption Assumption 2.4 $f(k_{i-1}, \omega_{it}, \xi_{it})$ is weakly increasing in ω_{it}, I conclude that $\Pr(i_{it} > z \mid k_{it-1}, \omega_{it})$ is weakly increasing in ω_{it}. It is well known that MLRP implies the first order stochastic dominance and ordering of expectations, so other results follow.

A.3 Proof of Proposition 5.2

Using Equation (3.3), I can write the moment function at the true parameter values as

$$m(w_{it}, \theta) = y_{it} - \theta_k k_{it} - \theta_l l_{it} = g(\omega_{it-1}) + \zeta_{it} + \eta^l_{it} + \epsilon_{it}.$$

Substituting $m(w_{it}, \theta) = g(\omega_{it-1}) + \zeta_{it} + \eta^l_{it} + \epsilon_{it}$ into Equation (3.5) we need to show that the following inequality holds

$$\mathbb{E}[g(\omega_{it-1}) + \zeta_{it} + \eta^l_{it} + \epsilon_{it} \mid k_{it}, i_{it-1} > z] - \mathbb{E}[g(\omega_{it-1}) + \zeta_{it} + \eta^l_{it} + \epsilon_{it} \mid k_{it}, i_{it-1} < z] \geqslant 0.$$

I proceed in two steps. First note that, by assumptions ζ_{it}, ϵ_{it} and ηit are orthogonal to information set at $t-1$. Therefore we have

$$\begin{aligned}\mathbb{E}[g(\omega_{it-1}) + \zeta_{it} + \epsilon_{it} + \eta_{it} \mid k_{it}, i_{it-1} > z] = \mathbb{E}[g(\omega_{it-1}) \mid k_{it}, i_{it-1} > z], \\ \mathbb{E}[g(\omega_{it-1}) + \zeta_{it} + \epsilon_{it} + \eta_{it} \mid k_{it}, i_{it-1} < z] = \mathbb{E}[g(\omega_{it-1}) \mid k_{it}, i_{it-1} < z].\end{aligned}$$

To finish the proof, I need to show that

$$\mathbb{E}[g(\omega_{it-1}) \mid k_{it}, i_{it-1} > z] \geqslant \mathbb{E}[g(\omega_{it-1}) \mid k_{it}, i_{it-1} < z].$$

To see this, by Proposition 2.1, I have a first order stochastic dominance for the distribution of ω_{it-1} conditional on 'high' and 'low' investment (conditional on k_{it-1}). Using this

$$F_{\omega_{it-1}}(t \mid k_{it}, i_{it-1} > z) \geqslant F_{\omega_{it-1}}(t \mid k_{it}, i_{it-1} < z), \qquad \text{for all } t > 0.$$

Since $g(\cdot)$ is a monotone function and stochastic order is preserved under the monotone transformation, we have

$$F_{g(\omega_{it-1})}(t \mid k_{it}, i_{it-1} > z) \geqslant F_{g(\omega_{it-1})}(t \mid k_{it}, i_{it-1} < z) \qquad \text{for all} \qquad t > 0.$$

This leads to the desired inequality

$$\mathbb{E}[g(\omega_{it-1}) \mid k_{it}, i_{it-1} > z] - \mathbb{E}[g(\omega_{it-1} \mid k_{it}, i_{it-1} < z] \geqslant 0.$$

B Other Moment Inequalities

B.1 Identification with Markov Assumption I

We have that

$$m(w_{it}, \theta) = y_{it} - \beta_k k_{it} - \beta_l l_{it} = \omega_{it} + \epsilon_{it} \tag{B.1}$$

$$m(w_{it-1}, \theta) = y_{it-1} - \beta_k k_{it-1} - \beta_l l_{it-1} = \omega_{it-1} + \epsilon_{it-1} \tag{B.2}$$

We have that

$$\mathbb{E}[m(w_{it}, \theta_0) m(w_{it-1}, \theta_0)] > 0 \tag{B.3}$$

B.2 Identification with Markov Assumption II

In this section, I show that the Markov property of ω_{it} in Assumption 2.1 provides moment inequalities and set identifies the production function. This result relies on the following proposition.

Proposition B.1. *Under Assumption 2.2 we have*

$$\mathbb{E}\Big[\big(\omega_{it} + \epsilon_{it} - \omega_{it-1} - \epsilon_{it-1}\big)^2\Big] \leqslant \mathbb{E}\Big[\big(\omega_{it} + \epsilon_{it} - \omega_{jt-1} - \epsilon_{jt-1}\big)^2\Big]. \tag{B.4}$$

Proof. See Appendix A.

This proposition states that the difference between productivity shocks across two periods is smaller for the same firm than for two different firms. The key assumption to obtain this result is that conditional distribution of ω_{it} is stochastically increasing in ω_{it-1}. Therefore, firm i's current period productivity, $\omega_{it} + \epsilon_{it}$, is closer to firm i's previous period productivity than firm j's previous period productivity in Euclidean distance. Now, using the Cobb-Douglas functional form, I write the productivity shocks in Proposition B.4 as

$$\Delta\omega_{it} + \Delta\epsilon_{it} = \Delta y_{it} - \theta_k \Delta k_{it} - \theta_l \Delta l_{it},$$
$$\Delta\omega_{ijt} + \Delta\epsilon_{ijt} = \Delta y_{ijt} - \theta_k \Delta k_{ijt} - \theta_l \Delta l_{ijt},$$

where I use $\Delta z_{it} := z_{it} - z_{it-1}$ and $\Delta z_{ijt} := z_{it} - z_{jt-1}$. Combining this with Proposition B.1, I construct a moment inequality

$$\mathbb{E}\Big[\Delta y_{it} - \theta_k \Delta k_{it} - \theta_l \Delta l_{it}\Big] \leqslant \mathbb{E}\Big[\Delta y_{ijt} - \theta_k \Delta k_{ijt} - \theta_l \Delta l_{ijt}\big)\Big], \tag{B.5}$$

which consists only of data and parameters, so it can be used for estimating bounds for the parameters. The result uses two assumptions: (i) Persistent productivity shock follows an exogenous Markov process, and (ii) Productivity shocks are additively separable in the production function. Therefore, moment inequalities are obtained under very general conditions. First, inputs could be dynamic or static. Second, we do not need to observe a proxy variable to control for productivity shocks. Finally, the variables that affect the firm's dynamic or static decisions are unrestricted. Of course, this flexibility might come with a cost, as the identified set might not be very informative.

In Proposition B.1, I use the Euclidean distance to derive moment inequalities. However, one can consider other distance measures and obtain different moment inequalities. In that case, different distances would give different identified sets, which can be intersected to obtain tighter bounds.

References

Ackerberg, Daniel A., Caves, Kevin, and Frazer, Garth. 2015. “Identification Properties of Recent Production Function Estimators”. *Econometrica* 83 (6): 2411–2451. (Cit. on p. 1).

Ackerberg, Daniel, Lanier Benkard, C., Berry, Steven, and Pakes, Ariel. 2007. “Chapter 63 Econometric Tools for Analyzing Market Outcomes”. In *Handbook of Econometrics*, 6:4171–4276. Elsevier. (Cit. on p. 1).

Andrews, Donald WK and Shi, Xiaoxia. 2013. “Inference Based on Conditional Moment Inequalities”. *Econometrica* 81 (2): 609–666. (Cit. on p. 19).

Balat, Jorge, Brambilla, Irene, and Sasaki, Yuya. 2019. “Heterogeneous Firms: Skilled-Labor Productivity and the Destination of Exports”. *Working Paper*. (Cit. on p. 4).

Blundell, Richard and Bond, Stephen. 2000. “GMM Estimation with Persistent Panel Data: An Application to Production Functions”. *Econometric Reviews* 19 (3): 321–340. (Cit. on p. 3).

Chernozhukov, Victor, Chetverikov, Denis, and Kato, Kengo. 2018. “Inference on Causal and Structural Parameters Using Many Moment Inequalities”. *The Review of Economic Studies* 86 (5): 1867–1900. (Cit. on pp. 19, 24).

Collard-Wexler, Allan and De Loecker, Jan. 2016. “Production Function Estimation with Measurement Error in Inputs”. *Working Paper*. (Cit. on pp. 3, 23).

Conley, Timothy G, Hansen, Christian B, and Rossi, Peter E. 2012. “Plausibly Exogenous”. *Review of Economics and Statistics* 94 (1): 260–272. (Cit. on p. 11).

Cooper, Russell W and Haltiwanger, John C. 2006. “On the Nature of Capital Adjustment Costs”. *The Review of Economic Studies* 73 (3): 611–633. (Cit. on p. 7).

Demirer, Mert. 2020. “Production function estimation with factor-augmenting technology: An application to markups”. *Job Market Paper*. (Cit. on p. 3).

Gandhi, Amit, Navarro, Salvador, and Rivers, David A. 2020. “On the Identification of Gross Output Production Functions”. *Journal of Political Economy* 128 (8): 2973–3016. (Cit. on p. 3).

Goolsbee, Austan and Gross, David B. 2000. “Estimating Adjustment Costs with Data on Heterogeneous Capital Goods”. *Working Paper*. (Cit. on p. 7).

Griliches, Zvi and Mairesse, Jacques. 1995. "Production Functions: the Search for Identification". *Working Paper 5067 National Bureau of Economic Research.* (Cit. on p. 3).

Hamermesh, Daniel S and Pfann, Gerard A. 1996. "Adjustment Costs in Factor Demand". *Journal of Economic Literature* 34 (3): 1264–1292. (Cit. on p. 7).

Hu, Yingyao, Huang, Guofang, and Sasaki, Yuya. 2011. "Estimating Production Functions with Robustness Against Errors in the Proxy Variables". *Working Paper.* (Cit. on pp. 3, 23).

Kasahara, Hiroyuki, Schrimpf, Paul, and Suzuki, Michio. 2015. "Identication and Estimation of Production Function with Unobserved Heterogeneity". *Working Paper.* (Cit. on p. 4).

Kim, Kyoo il, Petrin, Amil, and Song, Suyong. 2016. "Estimating Production Functions with Control Functions when Capital is Measured with Error". *Journal of Econometrics* 190 (2): 267–279. (Cit. on pp. 3, 23).

Levinsohn, James. 1993. "Testing the Imports-as-market-discipline Hypothesis". *Journal of International Economics* 35 (1): 1–22. (Cit. on p. 23).

Levinsohn, James and Petrin, Amil. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables". *The Review of Economic Studies* 70 (2): 317–341. (Cit. on pp. 1, 20).

Li, Tong and Sasaki, Yuya. 2017. "Constructive Identification of Heterogeneous Elasticities in the Cobb-Douglas Production Function". *arXiv:1711.10031 [econ].* (Cit. on p. 4).

Manski, Charles F. 1997. "Monotone Treatment Response". *Econometrica* 65 (6): 1311–1334. (Cit. on pp. 3, 10).

Manski, Charles F and Pepper, John V. 2000. "Monotone Instrumental Variables: With an Application to the Returns to Schooling". *Econometrica* 68 (4): 997–1010. (Cit. on pp. 3, 10).

Marschak, Jacob and Andrews, William H. 1944. "Random Simultaneous Equations and the Theory of Production". *Econometrica* 12 (3): 143–205. (Cit. on p. 3).

Mundlak, Yair. 1961. "Empirical Production Function Free of Management Bias". *Journal of Farm Economics* 43 (1): 44–56. (Cit. on p. 3).

Mundlak, Yair and Hoch, Irving. 1965. "Consequences of Alternative Specifications in Estimation of Cobb-Douglas Production Functions". *Econometrica* 33 (4): 814–828. (Cit. on p. 3).

Nevo, Aviv and Rosen, Adam M. 2012. "Identification with Imperfect Instruments". *Review of Economics and Statistics* 94 (3): 659–671. (Cit. on pp. 3, 11).

Olley, G. Steven and Pakes, Ariel. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry". *Econometrica* 64 (6): 1263. (Cit. on p. 1).

Semenova, Vira. 2017. "Machine Learning for Set-Identified Linear Models". *Working Paper, arXiv:1712.10024*. (Cit. on pp. 19, 24).

Shaked, Moshe and Shanthikumar, J George. 2007. *Stochastic Orders*. Springer Science & Business Media. (Cit. on pp. 14, 15).

Taymaz, Erol and Yilmaz, Kamil. 2015. "Estimating Plant-level Marginal Costs and Mark-ups in the Turkish Manufacturing Industry". *Middle East Technical University, ERC Working Paper*. (Cit. on p. 23).