Production Function Estimation Robust to Flexible Timing of Labor Input

Kyoo il Kim
Michigan State University

Yao Luo
University of Toronto

Yingjun Su*
IESR, Jinan University

March 2019

Abstract

Control function approaches to estimating production functions rely on a proxy that is monotone in an unobserved scalar productivity conditioning on other state variables. Ackerberg, Caves, and Frazer (2015) point out a potential functional dependence problem when conditioning only on capital. They provide a simple solution by conditioning on both capital and labor. This approach allows flexible timing of labor input for when firms learn all or part of their productivity. However, we demonstrate that ACF's moment condition may suffer from weak identification, and its significance depends on the timing of labor input. We then propose easy-to-implement modified procedures that remedy these issues, and provide Monte Carlo evidence. Estimation of production functions with flexible timing assumption of inputs is important because it can allow for unobserved, serially correlated, firm-specific wage shocks, labor being chosen prior to other variable inputs, or labor having dynamic implications. Moreover, since the exact timing of input choices is unknown and even may differ across firms in practice, our proposal is valuable as it fully incorporates the flexible framework but avoids weak identification.

Keywords: Production Function, Weak Identification, Control Function, Monte Carlo Simulation
JEL Classifications: C14, C18, D24

*Corresponding authors: Kyoo il Kim, kyookim@msu.edu; Yao Luo, yao.luo@utoronto.ca; Yingjun Su, yis17@pitt.edu

1 Introduction

The fundamental challenge in estimating production functions is due to the simultaneity of input and output decisions by firms. Since as early as Marschak and Andrews (1944), the literature has much focused on developing approaches to solving the problem. In the empirical literature, over the last two decades, the methods proposed by Olley and Pakes (OP, 1996) and Levinsohn and Petrin (LP, 2003) have been extensively used. Both OP and LP propose a control function approach to address this simultaneity by inverting unobserved productivity from a proxy variable. Their approach, however, relies critically on timing assumptions of input choices. In particular, LP assumes that labor input is determined after intermediate input so that the intermediate input demand does not depend on labor input. Also in their setting labor is a static input, and potential dynamic implications of labor input are all ruled out.

Ackerberg, Caves, and Frazer (ACF, 2015) argue that not only this timing assumption is restrictive in nature, but also a functional dependence problem arises in the first stage of the LP procedure that regresses output on labor input and a nonparametric function of other inputs.[1] As an alternative, they propose inverting an input demand function that is conditional on state variables, including labor input, which allows labor input to have potential dynamic implications. They argue that their approach is robust to different timing assumptions on labor input such that firms can choose the amount of labor input after learning about all or part of the unobserved productivity.

However, in this paper we first argue that ACF's moment condition for estimation may suffer from weak identification and that the degree to which it causes issues varies by the timing of labor input, and provide Monte Carlo evidence to illustrate our points. ACF attributes this identification problem only to a "spurious minimum." We show, however, that this spurious minimum arises only under a particular timing assumption: the current labor input is determined only after the current productivity shock is fully realized. Otherwise, this spurious minimum should not exist in the population. Nevertheless, our Monte Carlo simulations indicate that the spurious minimum is often found in finite samples unless the labor input is determined sufficiently ahead of the time at which the current productivity shock is realized. On the other hand, if the labor input is indeed determined far in advance of the current productivity shock, we argue that more serious weak identification problem arises in the ACF moments, and they may produce unreliable estimates. In this case, while one may augment current labor input to the instrument set, except that it is a valid instrument only when the labor input is determined before any part of the current productivity shock is realized.

[1] Wooldridge (2009) propose an approach to estimating the two stages in LP (or OP) simultaneously, avoiding this functional dependence issue. However, this approach does not relax the restrictive timing assumption of labor input as LP. See also Gandhi, Navarro, and Rivers (2017) for related identification problems of the proxy variable approach.

In this paper, we find easy-to-implement modifications that remedy the identification problem, with a resulting estimator that is robust to different timing of labor input. Specifically, we include further lagged inputs in the instrument set, as well as modify the ACF procedure by explicitly estimating the constant term in the production function. We also discuss why these modifications help for identification of production functions with flexible timing of labor input.

Monte Carlo simulations show that our modified procedure performs well in finite samples and yields estimates robust to different timing assumptions of labor input.[2] Moreover, in practice, we do not know when firms make their labor input choices, and different firms may have different timing of input choices. To demonstrate this point, we construct a data generating process (DGP) where the timing of input choices is heterogeneous across firms. Monte Carlo simulations demonstrate that our modified estimator exhibits robust performance in this setting as well. The results suggest our proposed estimator will be useful in the realistic situation where the exact timing of input choices is unknown and may differ across firms.

The remainder of this paper is organized as follows. Section 2 outlines production function estimations using control functions. Section 3 discusses the identification problems of ACF's moments. Section 4 exhibits our Monte Carlo experiments, in particular, with different timing of labor input. Section 5 develops our modified procedures and examines their finite sample performances. Section 6 concludes.

2 Review of Model and Estimation

In this section, we briefly review estimation of production functions using control functions. Since this paper develops estimation of production functions that allows flexible timing of labor input building on ACF, we focus on their approach in this review. Consider the value-added production function for a firm i at time t given by

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \eta_{it},$$

where l_{it} and k_{it} denote the natural log of labor and capital inputs, respectively; ω_{it} is the unobserved state variable that impacts the firm's decisions on inputs and production level; and η_{it} denotes a pure i.i.d. shock in production. The unobserved productivity ω_{it} is the source of the simultaneity, and the least squares estimation of the production function is inconsistent because the labor and capital input (or investment) choices depend on ω_{it}. k_{it} is a state variable, and l_{it} is a freely variable input in OP and LP while it is another state variable in ACF.

To control for this endogeneity problem, OP and LP propose using a proxy variable to invert

[2] Compared with this paper, the procedure proposed in Kim, Luo, and Su (2019) assumes this constant is known. Both ACF and Kim, Luo, and Su (2019) run their Monte Carlo studies only at a fixed timing of labor input, and do not examine the robustness of the estimator with different timing of labor input.

out unobserved productivity ω_{it} from the investment or input demand. Both OP and LP make an implicit assumption on the timing of labor input choice in their proxy variable approach. In particular LP assumes that labor input is determined after intermediate input, so that the intermediate input demand is given by

$$m_{it} = \tilde{f}_t(k_{it}, \omega_{it})$$

for which the capital input is the only observed state variable. ACF argues that this timing assumption can be restrictive, in addition to the functional dependence problem, since it does not allow l_{it} to be a state variable when the intermediate input is determined. To allow for more flexible assumptions on this timing of labor input choice, ACF defines the intermediate input demand as

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

and, given the monotonicity of the input demand in the unobserved scalar productivity, the inverse function

$$\omega_{it} = h_t(l_{it}, k_{it}, m_{it}) \tag{1}$$

becomes the control in their setting.[3] Specifically, the firm may choose labor input l_{it} at time $t - b$ where $b \in [0, 1]$. Note that this timing of labor input b is likely unknown to the analyst and heterogeneous across firms/industries. As a result, this specification allows labor to be determined before all or part of ω_{it} is realized at time t. Importantly, it allows labor to depend on serially correlated unobserved shocks such as wage, allows labor to have dynamic implications, that is, firm's decision on the current labor not only affects current profits, but also future profits, with possible adjustment costs, and allows intermediate inputs to be chosen after labor or simultaneously with labor. We build on this flexible setting to develop our modified procedures to estimate production functions later.

Substituting (1) into the production function, ACF obtains

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + h_t(l_{it}, k_{it}, m_{it}) + \eta_{it} = \Phi_t(l_{it}, k_{it}, m_{it}) + \eta_{it}. \tag{2}$$

In the first stage, ACF estimates the function $\Phi_t(l_{it}, k_{it}, m_{it})$ by minimizing the sample counterpart of the objective function $E[(y_{it} - \Phi_t(l_{it}, k_{it}, m_{it}))^2]$. In the second stage, ACF uses the

[3]Since m_{it} is a non-dynamic input, this monotonicity condition for m_{it}, being conditioned on both labor and capital (or only on capital in LP's setting), is identical to that for investment in OP such that more productive firms should use more intermediate inputs if they maximize their profits. The scalar productivity assumption rules out production functions with multiple structural unobservables e.g. as in McElroy (1987). However, the approaches with multiple errors typically utilize static first order conditions of input demands, so would rule out labor having dynamic implications, which we focus on in this paper.

estimated control function Φ_t and assumes a first order Markov process of the productivity as

$$\omega_{it} = E[\omega_{it}|\omega_{i,t-1}] + \xi_{it} = g(\omega_{i,t-1}) + \xi_{it}$$

to control for the endogeneity of inputs where ξ_{it} denotes an innovation term. Because the composite shock $\xi_{it} + \eta_{it}$ is uncorrelated with the firm's information at time $t-1$, ACF obtains the moment condition

$$\begin{aligned} E[\xi_{it} + \eta_{it}|I_{i,t-1}] = E[&y_{it} - \beta_0 - \beta_l l_{it} - \beta_k k_{it} \\ &- g\left(\Phi_{t-1}(l_{i,t-1}, k_{i,t-1}, m_{i,t-1}) - \beta_0 - \beta_l l_{i,t-1} - \beta_k k_{i,t-1}\right)|I_{i,t-1}] = 0, \end{aligned}$$

where $I_{i,t-1}$ denotes the firm's information at $t-1$.

Let $t-b$ denote a point in time when the firm chooses labor input l_{it} for $b \in [0,1]$. Depending on this timing assumption ACF proposes instruments $z_{it} = (1, k_{it}, l_{i,t-1}, \Phi_{t-1}(\cdot))$ if labor is to be chosen at time $t-b$ for any $b \in [0,1)$ and $z_{it} = (1, k_{it}, l_{it}, l_{i,t-1}, \Phi_{t-1}(\cdot))$ if labor is chosen at $t-1$, i.e. labor input depends on $\omega_{i,t-1}$, but not on any part of ξ_{it}. Here the fundamental identification argument, based on the timing assumptions, is that the state variable k_{it} is allowed to be correlated with $E[\omega_{it}|\omega_{i,t-1}]$, but it is uncorrelated with the innovation in the productivity shock $\xi_{it} = \omega_{it} - E[\omega_{it}|\omega_{i,t-1}]$, while only the lagged labor $l_{i,t-1}$ is uncorrelated with the innovation term ξ_{it} unless the current labor l_{it} is actually determined at time $t-1$ or before by the firm. On the other hand, both k_{it} and $l_{i,t-1}$ are uncorrelated with η_{it}, the pure i.i.d. shock in production at time t.

In practice, researchers often use a simple autoregressive model for the unobserved productivity as

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

and we focus on this simple setting. Let $w_{it} = \{(y_{it}, l_{it}) \cup z_{it}\}$ and write the moment condition:

$$\begin{aligned} E[r(w_{it};\theta_0)] &= 0 \qquad (3) \\ r(w_{it};\theta) &= z_{it} \times (y_{it} - \beta_0 - \beta_l l_{it} - \beta_k k_{it} \\ &\quad -\rho \cdot (\Phi_{t-1}(l_{i,t-1}, k_{i,t-1}, m_{i,t-1}) - \beta_0 - \beta_l l_{i,t-1} - \beta_k k_{i,t-1})), \end{aligned}$$

where $\theta = (\beta_0, \beta_l, \beta_k, \rho)$ and θ_0 denote the true parameter. Then, the production function parameters can be estimated using a GMM method based on the moment condition (3).

To facilitate the computation, ACF actually proposes to concentrate out two parameters, β_0 and ρ, similarly to LP method as follows. For trial values of the parameters β_l and β_k, first construct an estimate for $\beta_0 + \omega_{it}$ as

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

where $\widehat{\Phi}_t(l_{it}, k_{it}, m_{it})$ is obtained from the first stage. Then ACF runs regression $\beta_0 + \widehat{\omega_{it}(\beta_l, \beta_k)}$ on its lagged counterpart $\beta_0 + \widehat{\omega_{i,t-1}}(\beta_l, \beta_k)$ to obtain the regression residual $\hat{\xi}_{it}(\beta_l, \beta_k)$. Finally, ACF estimates β_l and β_k using the concentrated moment conditions

$$E\left[\hat{\xi}_{it}(\beta_l, \beta_k) \times \begin{pmatrix} l_{i,t-1} \\ k_{it} \end{pmatrix}\right] = 0 \tag{4}$$

if labor is chosen at time $t - b$ for any $b \in [0, 1)$; alternatively, ACF proposes

$$E\left[\hat{\xi}_{it}(\beta_l, \beta_k) \times \begin{pmatrix} l_{it} \\ l_{i,t-1} \\ k_{it} \end{pmatrix}\right] = 0 \tag{5}$$

if labor is chosen at time $t - 1$.[4]

In the next section we first argue that although these moments may allow for flexible timing of labor input, they are subject to potential identification problems. We then seek possible solutions in the subsequent section and develop modified procedures robust to the flexible timing framework.

3 Identification

In this section, we verify that the global identification problem arises in the ACF moments, but we also make the nuanced qualification that the problem should occur only for specific timing of labor input choice. On the other hand, we note that the moments (4) with the lagged labor input are potentially weak, even when the global identification problem does not exist. Additionally, while the moments (5) with current labor input are potentially strong, current labor is a valid instrument only when the labor is determined before any part of the current productivity shock is realized. These points are all illustrated in our Monte Carlo experiments later. We perform our identification analysis based on the DGP setting considered by ACF, which is standard in the empirical literature of production function estimations except that we consider flexible timing of labor input by firms.[5]

[4]Note, however, that to use the alternative moment (5), one needs to know the exact timing of labor input, and the timing should be the same for all firms.

[5]Analogously to the current setting, our identification analysis here may be applied to other settings (e.g. different processes of productivity change, wages, or other labor adjustment costs) for which we may find DGPs

Following ACF, we let productivity $\omega_{i,t-1}$ evolve to $\omega_{i,t-b}$, where $b \in [0,1]$, at which point in time the firm chooses labor input l_{it}. Specifically,

$$\omega_{i,t-b} = \rho^{1-b}\omega_{i,t-1} + \xi_{it}^A,$$
$$\omega_{it} = \rho^b \omega_{i,t-b} + \xi_{it}^B,$$

where it is imposed that $Var(\rho^b\xi_{it}^A + \xi_{it}^B) = Var(\xi_{it})$, and hence $Var(\xi_{it}^A)$ is decreasing in b while $Var(\xi_{it}^B)$ is increasing in b.[6] In this setting, firms are allowed to choose l_{it} with less than perfect information about ω_{it}, and this information decreases as b increases. In particular, $\xi_{it}^B = 0$ and $Var(\xi_{it}^B) \equiv \sigma_b^2 = 0$ when $b = 0$. Therefore, this framework allows labor input l_{it}, chosen at period t, period $t-1$, or period $t-b$ (with $0 < b < 1$), to have potential dynamic implications.

To see how a spurious identification point arises in the original ACF moments, let us first examine the inverted productivity at the true parameters. From (2), by construction of $\Phi_t(\cdot)$, we have

$$\Phi_t(l_{it}, k_{it}, m_{it}) - \beta_k k_{it} - \beta_l l_{it} = \beta_0 + \omega_{it}. \tag{6}$$

In the ACF concentrated procedure, $\widehat{\beta_0 + \omega_{it}} = \widehat{\Phi}_t(l_{it}, k_{it}, m_{it}) - \beta_k k_{it} - \beta_l l_{it}$ is used for the autoregressive regression of the productivity in place of ω_{it}. Specifically, they estimate

$$(\widehat{\beta_0 + \omega_{it}}) = \alpha_0 + \rho(\widehat{\beta_0 + \omega_{i,t-1}}) + \xi_{it}, \tag{7}$$

and the above regression becomes equivalent to $\widehat{\omega}_{it} = \rho\widehat{\omega}_{i,t-1} + \xi_{it}$ with $\alpha_0 = \beta_0(1-\rho)$. Note, however, that including an intercept in the regression makes the residual $\xi_{it}(\beta_l, \beta_k)$ have mean zero by construction, regardless of whether the intercept is actually equal to the true $\alpha_0 = \beta_0(1-\rho)$. We later discuss why this fact becomes important for our proposed procedures with flexible timing of labor input.

3.1 The "spurious" minimum problem

To see how identification could fail in the original ACF moments, as noted by ACF, and to further elaborate on the issue, let the spurious minimum be $\tilde{\beta}_k = 0$ and $\tilde{\beta}_l = \beta_l + \beta_k = 1$. Let us write $\Phi_{it} = \Phi_t(l_{it}, k_{it}, m_{it})$ for ease of notation. From (6), it follows that

$$\Phi_{it} - \tilde{\beta}_k k_{it} - \tilde{\beta}_l l_{it} = \beta_0 + \beta_k k_{it} - (1-\beta_l) l_{it} + \omega_{it}. \tag{8}$$

under which similar identification problems exist. We focus on the current setting since ACF develops their estimation procedure for this standard setting in the empirical literature.

[6] See ACF's DGP in their Monte Carlo experiments, and discussions following their equation (37).

Also, from the optimal labor input chosen by firms maximizing expected profits at time $t-b$, given the production function and the firm-specific wage, we can write $(1-\beta_l)l_{it}$ as

$$(1-\beta_l)l_{it} = \beta_0 + \ln\beta_l - \ln W_{it} + \beta_k k_{it} + (\rho^b \omega_{i,t-b} + \frac{1}{2}\sigma_b^2), \tag{9}$$

where W_{it} denotes the wage, and the productivity $\omega_{i,t-1}$ evolves to $\omega_{i,t-b}$, at which point in time the firm chooses labor input.[7] Plugging (9) into (8), we obtain

$$\Phi_{it} - \tilde{\beta}_k k_{it} - \tilde{\beta}_l l_{it} = -\ln\beta_l - \sigma_b^2/2 + \ln W_{it} - \rho^b \omega_{i,t-b} + \omega_{it} = (-\ln\beta_l - \sigma_b^2/2) + \xi_{it}^B + \ln W_{it}, \tag{10}$$

because $\omega_{it} = \rho^b \omega_{i,t-b} + \xi_{it}^B$.

Equation (10) has two important implications for the identification problem. First, the constant term $(-\ln\beta_l - \sigma_b^2/2)$ on the RHS of (10) is different from β_0 in (6). We utilize this fact to modify the ACF procedure to later aid identification. Second, the regression residual from the AR(1) regression at the spurious minimum using $\Phi_{it} - \tilde{\beta}_k k_{it} - \tilde{\beta}_l l_{it}$ is also different from ξ_{it} in (7). Now suppose the wage is serially correlated, and follows an AR(1) process as

$$\ln W_{it} = \rho_W \ln W_{i,t-1} + \xi_{it}^W,$$

where ξ_{it}^W denotes the current period wage shock, is not correlated with firm's information available at time $t-1$. From (10), we then obtain

$$\Phi_{it} - \tilde{\beta}_k k_{it} - \tilde{\beta}_l l_{it} - \rho_W \left(\Phi_{i,t-1} - \tilde{\beta}_k k_{i,t-1} - \tilde{\beta}_l l_{i,t-1}\right) + \tilde{\alpha} = \xi_{it}^B - \rho_W \xi_{i,t-1}^B + \xi_{it}^W, \tag{11}$$

where $\tilde{\alpha} = (1-\rho_W)(\ln\beta_l + \sigma_b^2/2)$.

Note that innovation term ξ_{it}^W in the wage process is independent of $(k_{it}, l_{i,t-1})$ because they are determined at time $t-1$. Therefore, as discussed in ACF's footnote 16, the spurious parameter $(\tilde{\beta}_k = 0, \tilde{\beta}_l = 1)$ may solve the ACF moment condition. However, this is true only at an extreme timing of labor choice because the composite error in the spurious residual function (11), $\xi_{it}^B - \rho_W \xi_{i,t-1}^B + \xi_{it}^W$, satisfies the moment condition with the ACF instruments $(k_{it}, l_{i,t-1})$ only if $b = 0$. In fact, while $l_{i,t-1}$ satisfies the spurious moment condition for all $b \in [0,1]$, k_{it} is a function of lagged productivity $\omega_{i,t-1}$, and is correlated with $\xi_{i,t-1}^B$, unless $b = 0$. This is because $\xi_{i,t-1}^B$ is part of $\xi_{i,t-1}$, and hence is correlated with $\omega_{i,t-1}$ and, by extension, with k_{it}. From these observations, we conclude $l_{i,t-1}$ is potentially susceptible to the identification problem for all $b \in [0,1]$, while k_{it} becomes problematic only if b approaches zero.

In summary, the spurious solution of the ACF moments in the population arises only when $b = 0$, in which case $\xi_{it}^B = \xi_{i,t-1}^B = 0$ in (11). In the population, as long as $b > 0$, we see

[7] Therefore, ACF's DGP assumes that W_{it} is known to firms at time $t-b$.

that the spurious minimum should not arise because the instrument k_{it} is correlated with $\xi^B_{i,t-1}$, violating the spurious moment condition.

3.2 Weak identification

In this subsection, we discuss why the ACF moments are subject to weak identification even when b is away from zero. Note that b is set to 0.5 in the original ACF DGP, so there should be no spurious minimum in the population of this DGP as discussed in the previous section. However, in finite samples, the spurious minimum problem may exist if the correlation between $\xi^B_{i,t-1}$ and the instrument k_{it} is small and/or b is close to zero. In fact, our Monte Carlo experiments suggest that the spurious minimum arises for various timing of labor choice (see also Kim, Luo, and Su (2019)'s extensive Monte Carlo experiments using the original ACF DGP with $b = 0.5$). From our discussion above, $l_{i,t-1}$ is mean independent of the spurious residual in (11) for all $b \in [0, 1]$. Therefore, although $l_{i,t-1}$ is a valid instrument regardless of the timing of labor choice $b \in [0, 1]$, it is potentially weak, yielding spurious solutions in finite samples.

Another way to see this weak instrument problem is from a regression perspective of moment condition (3). Given a value of ρ, $l_{it} - \rho l_{i,t-1}$ can be regarded as the regressor for the labor coefficient. Thus, if the correlation between instrument $l_{i,t-1}$ and regressor $l_{it} - \rho l_{i,t-1}$ is not sufficiently significant for some ρ, then instrument $l_{i,t-1}$ becomes weak. In other words, the differencing $l_{it} - \rho l_{i,t-1}$ may effectively dispense with the *signal* in the regressor and the *noise* becomes dominant. This argument is similar to the panel data approach to estimating production functions with a fixed effect (see, e.g., Blundell and Bond (2000)). From the labor input function in (9), we see that the only remaining exogenous variation of labor input, net of capital input and lagged productivity, is the wage, since capital input and lagged productivity also appear in the residual function (3). Therefore, the differencing $l_{it} - \rho l_{i,t-1}$ at or near spurious autoregressive coefficient ρ_W (i.e., the autoregressive coefficient of the wage process) will weaken the correlation between $l_{it} - \rho l_{i,t-1}$ and instrument $l_{i,t-1}$, which may yield the weak instrument problem.[8]

On the other hand, current period labor choice l_{it} is not subject to the spurious minimum issue because it is always correlated with ξ^W_{it}, i.e., $E[\xi^W_{it} l_{it}] \neq 0$, in (11) regardless of the timing b. However, l_{it} is not a valid instrument unless $b = 1$, because $E[\xi_{it} l_{it}] \neq 0$ as long as $b \neq 1$. Nevertheless, its invalidity is less severe when b approaches 1, such that labor input is determined at a point in time close to the previous period $t - 1$ (when the firm has very little

[8]We also note that ACF estimates both β_l and β_k in the second step. Since labor l_{it} and capital k_{it} are both endogenous variables, identification requires at least two instrumental variables. ACF moments use $l_{i,t-1}$ and k_{it} for instruments. However, significant correlation among multiple instruments can reduce estimation precision in finite samples, due to the potential violation of a proper rank condition for identification. In fact, both instruments in ACF moments are functions of lagged wage $\ln W_{i,t-1}$ and lagged capital $k_{i,t-1}$ and thus are subject to some non-trivial correlation.

information on the current productivity shock other than the previous period's $\omega_{i,t-1}$).

In summary, we find that $l_{i,t-1}$ is a valid instrument regardless of labor input timing $b \in [0, 1]$, but it is potentially a weak instrument, which may yield spurious solutions in finite samples. On the other hand, l_{it} is a strong instrument, but invalid unless $b = 1$. In other words, the ACF procedure with the moments (4) may suffer from global identification failure if b approaches zero and suffer from the weak identification issue if b gets close to one. On the other hand, the ACF procedure with the moments (5) does not suffer from either problem, but it produces inconsistent estimates unless $b = 1$. The first part of our Monte Carlo study in Section 4 clearly illustrates these points.

Note that ACF's general framework, using the conditional input demand function, allows flexible timing of labor input for when the firm learns all or part of the productivity. Therefore, it is valuable to develop an estimation procedure that fully incorporates the flexible framework but avoids weak identification. The second part of our Monte Carlo study in Section 5 illustrates such procedures we develop in this paper.

4 Monte Carlo Study of the ACF Procedure

Our design of the experiments in Section 4 and 5 is twofold. First, we conduct experiments to demonstrate the identification issues of the original ACF procedure, as discussed in Section 3. We estimate the production function by replicating the original ACF procedure but varying the initial value for estimation, searching for the global minimum of the GMM objective function.[9] We follow the Monte Carlo setup (DGP1) employed in ACF to generate the data, except that we vary the timing of labor input choice such that firms are allowed to choose their labor inputs at $t - b$ where $b \in \{0, 0.05, 0.1, ..., 0.95, 1\}$. Note that the original ACF DGP assumes $b = 0.5$. Second, we propose modified procedures that overcome weak identification, and are robust to flexible timing of labor input. We then demonstrate their performance through simulations in Section 5. There, we also construct a DGP in which the timing of labor input is heterogeneous across firms and find that our proposed estimator is robust to heterogeneous timing of labor input choice.

In all exercises, we try different starting values of (β_l, β_k), i.e. (0, 1), (0.1, 0.9), (0.2, 0.8),...., (0.9, 0.1), for the search of the global minimum in the estimation. In particular, for each pair of initial values, we obtain an estimate of (β_l, β_k) that minimizes the objective function (i.e., local minimum). We then compare the values of these local minima to obtain our estimate of the global minimum for each simulated dataset. Moreover, for our estimation with the over-identified moment conditions, we use the continuously updated GMM estimator (CUE) (see,

[9] ACF uses the true value as the initial value to set aside the identification caveat in their Monte Carlo study. See their footnote 16.

e.g., Hansen, Heaton, and Yaron (1996)). Lastly, we report the empirical distributions of the estimates in box plots by the timing of labor input choice described in the main text, and also tabulate the means and standard deviations in the Appendix.

4.1 ACF Procedure with different timing of labor input

4.1.1 The identification problem

As discussed in Section 3, the ACF procedure may suffer from identification problems depending on the timing of labor input. We construct our first experiment to measure how this "spurious" minimum problem or weak identification influences estimation with finite samples under different timing assumptions of labor input choice. For a given timing of labor input $b \in \{0, 0.05, 0.1, ..., 0.95, 1\}$, we simulate 1,000 datasets. Each dataset contains 1,000 firms for 10 time periods. We estimate the production function by following the standard ACF procedure. That is, we adopt the concentrated moment condition (4) with instruments $(l_{i,t-1}, k_{it})$ to estimate β_l and β_k.

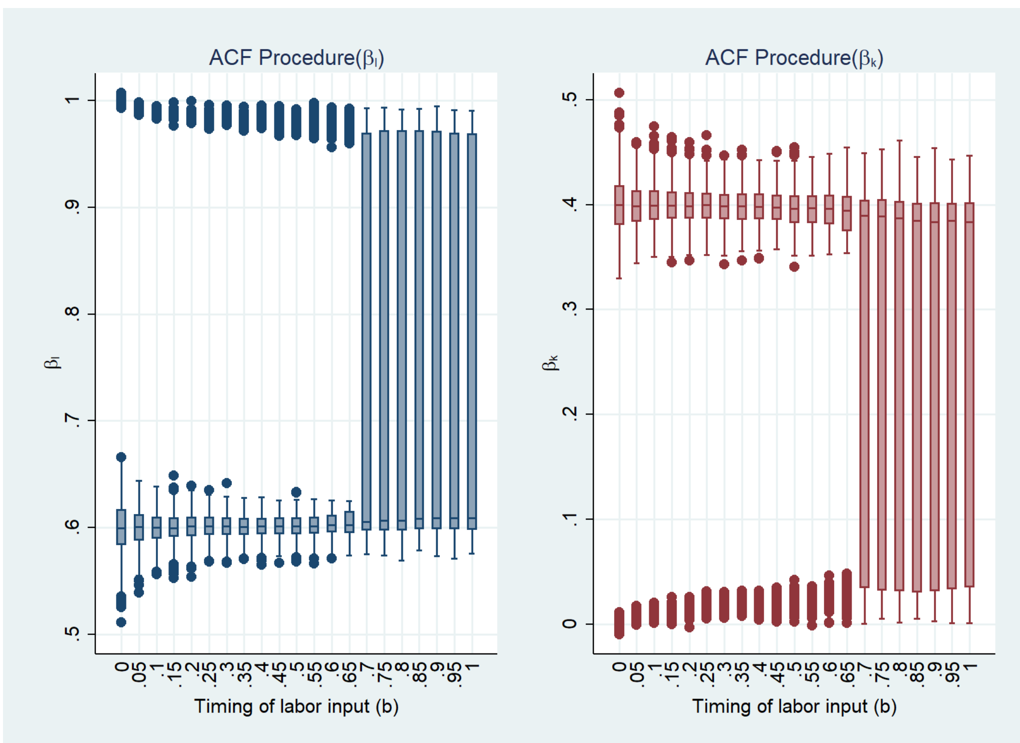

Figure 1: ACF Procedure with Different Timing of Labor Input

Note: The box plots are generated using estimates from 1,000 simulated datasets for each timing of labor input. For each value of b, each firm is assumed to choose its current labor input l_{it} at $t-b$. The true parameter values are $(\beta_l, \beta_k) = (0.6, 0.4)$.

Figure 1 clearly illustrates that the "spurious" minimum or weak identification exists and that the degree to which the identification problem occurs depends on different timing assumptions of labor input choice.[10] Specifically, these box plots indicate that when $b < 0.7$, the spurious minimum estimates around $(\beta_l, \beta_k) = (1, 0)$ arise more frequently. Note that the true parameter values are $(\beta_l, \beta_k) = (0.6, 0.4)$ for all b in the DGP. On the other hand, when $b \geq 0.7$, the ACF procedure yields estimates that are more dispersed, measured by the first and the third quartile of the simulated estimates, suggesting weak identification of the ACF moment condition (4).

In the Appendix, Table 1 reports averages and standard deviations of the estimated coefficients that yield the global minimum from this procedure with 1,000 repetitions. Table 1 further illustrates the identification problem of the ACF procedure showing that the mean of the estimated β_l increases from 0.610 to 0.749 and the mean of β_k decreases from 0.389 to 0.251 as b increases from 0 to 1. Both patterns point to the fact that the estimated coefficients deviate further from the true values as labor has more dynamic features. In other words, the ACF procedure suffers more severely from the identification problem as labor input is determined further in advance from the current period t, at which point in time the current productivity is fully learned by the firm. Also, the standard deviations of the 1,000 estimates become bigger as b increases, implying that the ACF estimator suffers more greatly from weak identification as labor is chosen closer to the previous period $t - 1$.

4.1.2 l_{it} as an additional instrument

ACF remarks that, if labor is chosen at $t - 1$, one can use l_{it} as an additional instrument. We now demonstrate the performance of such adjustment under different timing of labor choice. Specifically, we adopt the concentrated moment conditions in equation (5) with instruments $(l_{it}, l_{i,t-1}, k_{it})$ to estimate β_l and β_k.

Figure 2 clearly shows that, while this approach does not produce a spurious minimum, compared to the estimator using moment condition (4), this approach produces biased estimates of β_l and β_k unless $b = 1$, i.e., when labor is chosen at time $t - 1$. In fact, when $b \geq 0.65$, the estimates tend to converge to the true values faster as b increases. For example, Table 2 shows that the mean of the estimated β_l drops from 0.782 to the true value 0.6 as b increases from 0.65 to 1, whereas the mean of the estimated β_l only slightly drops from 0.944 to 0.884 as b increases from 0 to 0.6. This experiment confirms that while l_{it} is a more relevant IV and it is not subject to a spurious minimum problem, it is valid only when $b = 1$.

[10] This incidence of the spurious minimum or weak identification may substantially vary by different DGPs. In the original ACF DGP we use here the standard deviation of the log wage is set to be constant over time and equal to 0.1. Kim, Luo, and Su (2019) experiment with an alternative DGP (with fixed $b = 0.5$) for which the standard deviation of log wage is increased to 0.5, and find that the percentage of spurious minimum indications, e.g. for cases the estimated $\beta_l > 0.80$ and $\beta_k < 0.20$, increases from 8.5% to 39.1% with the alternative DGP of a higher standard deviation of log wage.

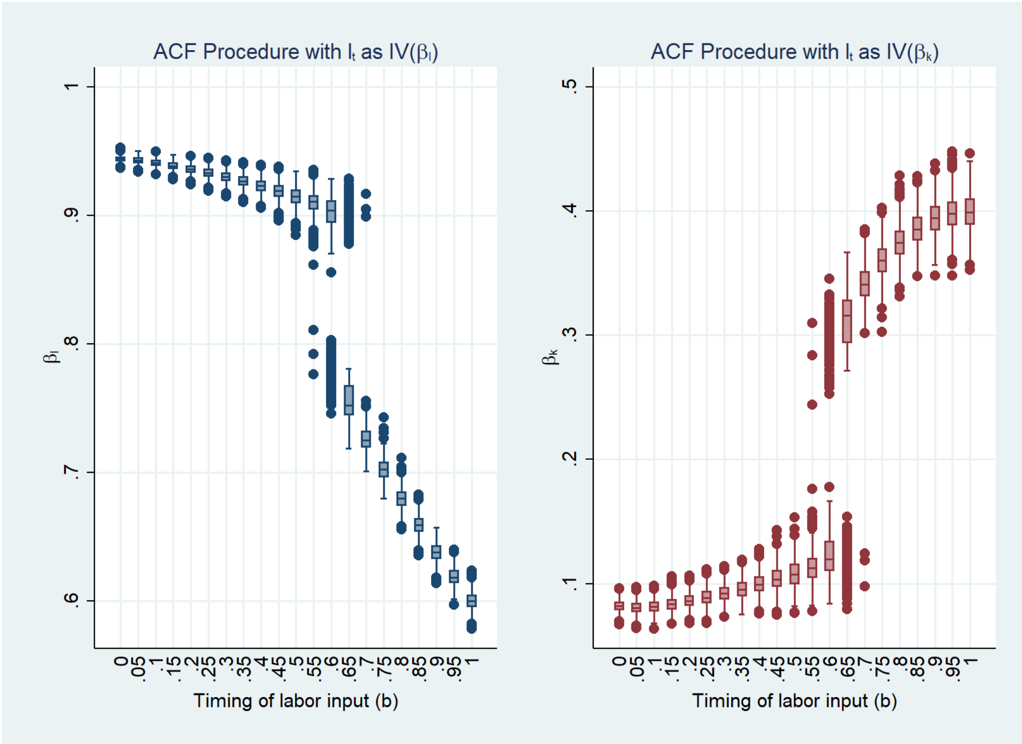

Figure 2: ACF Procedure with l_{it} as IV

In summary, we conclude that the spurious minimum problem or the weak identification of the ACF procedure with moment (4) is not negligible in finite samples. The identification problem prevails under all timing assumptions of labor input, and severity of the problem increases when labor is determined at a point in time closer to $t-1$. On the other hand, even though l_{it} is a more relevant IV in alternative moment condition (5), it is only valid when $b = 1$.

5 Modified Procedures Robust to Flexible Timing of Labor

The goal of this section is to develop procedures that are robust to different timing of labor input. We also illustrate the finite sample performance of our proposed estimators using the same Monte Carlo setup from Section 4.

We make three important modifications to the original ACF procedure: (i) We run the AR regression of productivity without an intercept. This may require us to identify the constant term β_0 in the production function. In our first modified procedure, we assume β_0 is known to the econometrician, which helps us to gain some intuition about identification. In the subsequent sections, we generalize our procedure to estimate this constant term in an earlier step or along with other parameters. (ii) We also add "1" to the instrument set. This ensures that the AR regression's residual $\hat{\xi}_{it}(\beta_l, \beta_k)$ has zero sample mean, which is the sample analogue of the innovation term ξ_{it} having zero mean.[11] (iii) We add further lagged input variables such as $l_{i,t-2}$ and $k_{i,t-1}$ to the instrument set (4). The further lagged input variables act as excluded instruments, and they are relevant instruments as illustrated in Kim, Luo and Su (2019).

In the proxy variable approach for estimating production functions, it is common practice to ignore various other unobservable factors of input demand, which may indicate market/industry structure, input prices (Olley and Pakes, 1996), or other aggregate shocks (Hahn et al, 2017). In the ACF setting, this missing factor of the input demand function is the wage of the individual firm. Since the wage follows an autoregressive process and is a factor of the labor input demand, lagged labor input $l_{i,t-2}$ becomes a relevant instrument for l_{it}. Also, the further lagged labor $l_{i,t-2}$ is not as highly correlated to k_{it} as $l_{i,t-1}$ is to k_{it}, thereby bringing additional exogenous variation to the instrument set to aid identification.

If we had included all the factors that determine the input demands, including firm-specific input prices, in $m_{it} = f_t(l_{it}, k_{it}, \omega_{it})$ or in its inverse $\omega_{it} = f_t^{-1}(l_{it}, k_{it}, m_{it})$ and so the control function $\Phi_t(l_{it}, k_{it}, m_{it})$, the further lagged labor inputs would not be relevant or less relevant for current labor.[12] To illustrate this point, let the t subscript in the input demands denote input prices, and let $l_{it} = q(k_{it}, \omega_{i,t}, p_{it})$ be the labor input, $m_{it} = f(l_{it}, k_{it}, \omega_{it}, p_{it})$ be the intermediate input where p_{it} denotes a vector of firm-specific input prices including wages, and note that

$$\begin{aligned} l_{it} &= q(k_{it}, f^{-1}(l_{it}, k_{it}, m_{it}, p_{it}), p_{it}) \\ &= q(k_{it}, g(f^{-1}(l_{i,t-1}, k_{i,t-1}, m_{i,t-1}, p_{i,t-1})) + \xi_{it}, p_{it}) \\ &= q(k_{it}, g(\Phi(l_{i,t-1}, k_{i,t-1}, m_{i,t-1}, p_{i,t-1}) - \beta_0 - \beta_l l_{i,t-1} - \beta_k k_{i,t-1}) + \xi_{it}, p_{it}), \end{aligned}$$

where $\omega_{it} = g(\omega_{i,t-1}) + \xi_{it}$ follows the first-order Markov process. Therefore, $l_{i,t-2}$ should not be relevant for current labor input given $(k_{it}, k_{i,t-1}, l_{i,t-1}, m_{i,t-1}, p_{it}, p_{i,t-1})$. However, $l_{i,t-2}$ remains relevant because in the current empirical setting the firm-specific input prices including wages are not observed to econometricians, so are not included in the input demands, and the control function for estimation.[13]

5.1 Estimation with known β_0

We discuss why the first modification helps for identification. If we run the AR regression without the intercept, it helps the moment condition to yield a "true" solution, deterring the

[11]Note that since we run the AR regression without an intercept, the regression residual does not necessarily have zero mean, and imposing the zero mean condition acts like an additional moment condition.

[12]It is generally not possible in the estimation of production functions to include all the factors in the input demand due to limitations of available data (e.g. individual firm's wages or other input prices are not observed).

[13]Kim, Petrin, and Song (2016) consider the input demand functions as $l_{it} = q(k_{it}, \omega_{it}, p_{it})$ and $m_{it} = f(k_{it}, \omega_{it}, p_{it})$ in the LP setting, and show that the current input prices can be used as instruments for l_{it} even when the labor and the intermediate input depend on the same state variables.

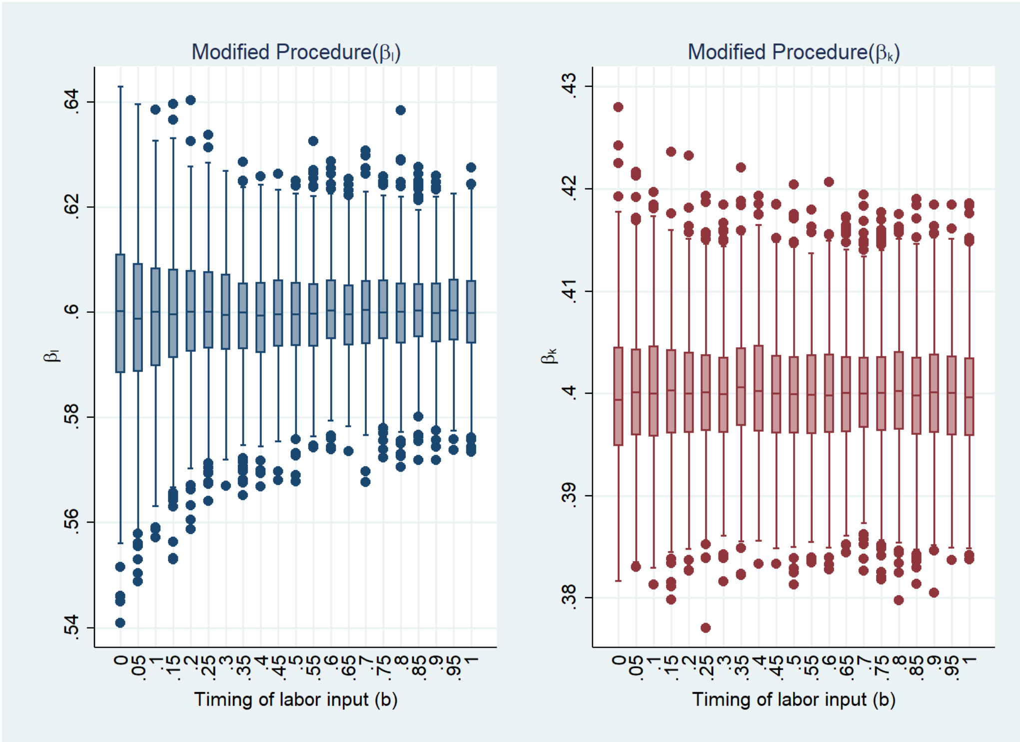

Figure 3: Modified Procedure with Augmented IV (Known β_0)

spurious minimum, because the spurious minimum solution requires the AR regression in (11) to have a non-zero intercept as $\tilde{\alpha} = (1 - \rho_W)(\ln \beta_l + \sigma_b^2/2)$, which is different from the true intercept $\alpha_0 = \beta_0(1 - \rho)$ in (7). In the ACF DGP, we have $\beta_0 = 0$ in the production function, allowing distinction between true intercept $\alpha_0 = 0$ and spurious intercept $\tilde{\alpha}$. Indeed, this argument holds for any β_0, since we can always subtract β_0 out from $\widehat{\beta_0 + \omega_{it}}$ in AR regression (7), if β_0 is known. Without the intercept in the "spurious" AR regression using (11), the "spurious" innovation term $\xi_{it}^B - \rho_W \xi_{it}^B + \xi_{it}^W$ would not have zero mean, which is inconsistent with the true DGP, in which the true innovation term ξ_{it} does have zero mean.

In the modified procedure, the AR regression does not include the intercept; i.e., we run the regression $\widehat{\omega}_{it} = \rho \widehat{\omega}_{i,t-1} + \xi_{it}$ to obtain the residual. This regression without an intercept deters ξ_{it} from having zero mean at parameter values other than the true one. Instead, we include "1" in the instrument set so that we can ensure innovation ξ_{it} has zero mean at the true parameter. If we included a constant in the regression, the residual would have zero mean for any parameter values that include the spurious solution.

Figure 3 and Table 3 report the estimation results, using the same simulated data set as in Section 4, from the modified procedure with the augmented instruments $(1, l_{i,t-1}, l_{i,t-2}, k_{it}, k_{i,t-1})$ assuming β_0 is known. We use the CUE GMM estimator and vary the initial value of estimation, searching for the global minimum of the objective function, as in Section 4. The modified procedure obviously improves the estimates of the production function parameters. The means of the estimated coefficients are very close to the truth and the standard deviations of the 1,000 replicated estimates are also much smaller than the estimates from the original ACF procedure (see Figure 1 and Table 1), indicating that virtually all estimates from the modified procedure are driven by the true solution of the population moment condition.

Importantly, the modified procedure is robust to different timing of labor input, while the original ACF procedure is somewhat sensitive to the DGPs with different timing of labor input. Also, the standard deviations of the 1,000 estimates for the DGPs with higher values of b are slightly smaller than those of the DGPs with smaller values of b, while the original ACF procedure suffered from more severe weak identification when b approaches one. Indeed, the standard deviation of the estimates from the modified procedure increases marginally as b approaches zero, at the point where the original ACF moments yield the spurious minimum in the population. Nevertheless, the modified procedure does not yield a spurious minimum, even at $b = 0$.

5.2 Estimating β_0 in the first step

In practice, the constant in the production function is unknown. Moreover, it may not be separately identified from the mean of the productivity. To deal with this issue, we first estimate β_0 in the first stage of the ACF procedure. This constant β_0 is interpreted as either the sum of the constant in the production function and the mean of the productivity, or just the constant in the production function if one normalizes the mean of the productivity to zero. Note that, from (6) we can obtain this constant from the estimate of the constant term in $\widehat{\Phi}_t(l_{it}, k_{it}, m_{it})$.

Let the constant term in $\widehat{\Phi}_{it}$ be $\hat{\phi}_0$. Removing this constant from $\widehat{\beta_0 + \omega_{it}} = \widehat{\Phi}_t(l_{it}, k_{it}, m_{it}) - \beta_k k_{it} - \beta_l l_{it}$, we then obtain the de-meaned productivity $\widetilde{\omega}_{it} = \widehat{\beta_0 + \omega_{it}} - \hat{\phi}_0$. Lastly, we can use regression $\widetilde{\omega}_{it} = \rho \widetilde{\omega}_{i,t-1} + \xi_{it}$ to obtain the innovation term as if the productivity has zero mean. We then estimate the production function parameters using the CUE GMM estimator with augmented instruments $(1, l_{i,t-1}, l_{i,t-2}, k_{it}, k_{i,t-1})$.

Alternatively, one can derive function $\Phi_t(\cdot)$, given a specific production function, and obtain the constant term. Consider a Leontief production function, as in ACF,

$$Y_{it} = \min \left\{ e^{\beta_0} K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it}}, e^{\beta_m} M_{it} \right\} e^{\eta_{it}},$$

where $Y_{it} = e^{y_{it}}$, $K_{it} = e^{k_{it}}$, $L_{it} = e^{l_{it}}$, and $M_{it} = e^{m_{it}}$. In this Leontief production function setting, we have

$$\beta_m + m_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it}, \tag{12}$$

and hence substituting ω_{it} — the inverse intermediate input demand function — into the value-

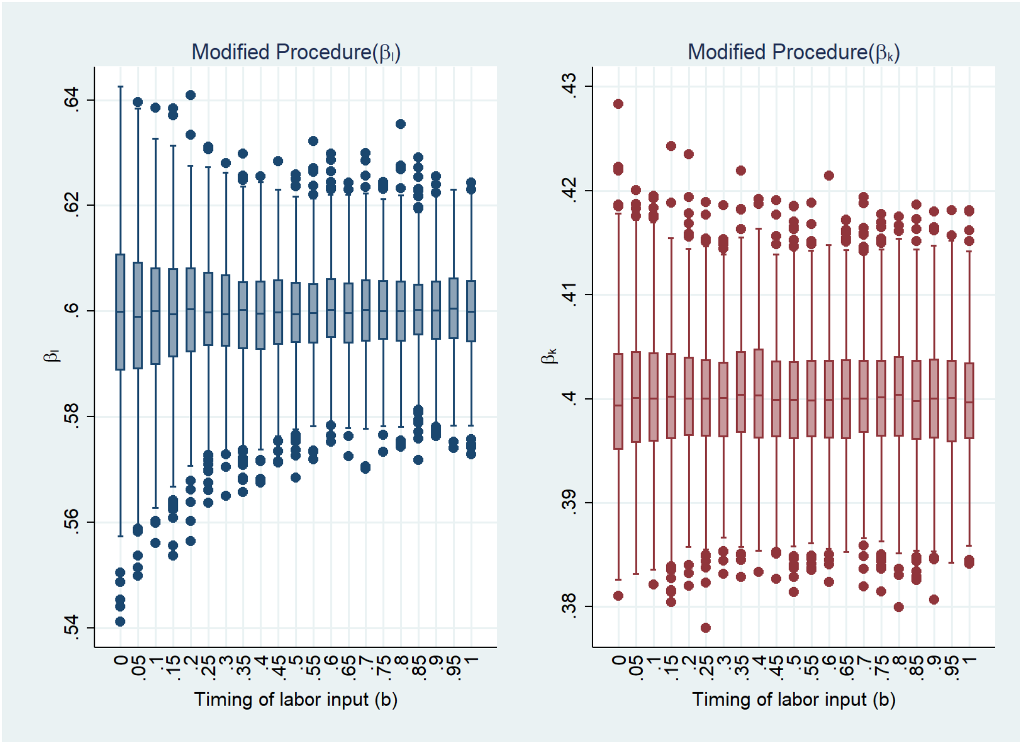

Figure 4: Modified Procedure with Augmented IV (Estimating β_0 in the first step)

added production function, we obtain

$$\begin{aligned} y_{it} &= \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it} \\ &= \beta_0 + \beta_k k_{it} + \beta_l l_{it} + [\beta_m + m_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it}] + \eta_{it} \\ &= \beta_m + m_{it} + \eta_{it}. \end{aligned} \tag{13}$$

This implies the true function $\Phi_t(l_{it}, k_{it}, m_{it})$ of the ACF DGP first stage becomes $\Phi_t(l_{it}, k_{it}, m_{it}) = \beta_m + m_{it}$. Then, following equation (13), we can estimate constant β_m using OLS of y_{it} on m_{it}. Moreover, from equation (12), it is clear that β_m equals the sum of the constant in the production function and the mean of the productivity in this specification.

Figure 4 and Table 4 report the results, which suggest that our modified procedure is still robust to different timing of labor input, even after we deal with unknown constant β_0. Both the means of the estimated β_l and β_k are very close to the true values, and the standard deviations of the 1,000 estimates are much smaller than those from the original ACF procedure. Interestingly, we observe that the performance of the modified procedure estimating the unknown constant is almost identical to the one in the previous experiment with known β_0.

5.3 Estimating β_0 jointly with (β_l, β_k)

A more general approach we propose is to estimate $(\beta_0, \beta_l, \beta_k)$ simultaneously. For this purpose, we modify the concentrated moment condition as follows.

1. For a trial value of the parameters $(\beta_0, \beta_l, \beta_k)$, first construct an estimate for ω_{it} as

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

   where $\widehat{\Phi}_t(l_{it}, k_{it}, m_{it})$ is obtained from the ACF first stage, and run the AR(1) regression without a constant as

$$\omega_{it}(\widehat{\beta_0, \beta_l}, \beta_k) = \rho \times \omega_{i,t-1}(\widehat{\beta_0, \beta_l}, \beta_k) + \xi_{it}(\beta_0, \beta_l, \beta_k)$$

   to obtain the regression residual $\hat{\xi}_{it}(\beta_0, \beta_l, \beta_k)$.

   Extending this step to a first-order Markov process of ω_{it} is also straightforward. In this case we estimate the regression

$$\omega_{it}(\widehat{\beta_0, \beta_l}, \beta_k) = g(\omega_{i,t-1}(\widehat{\beta_0, \beta_l}, \beta_k); \beta_\omega) + \xi_{it}(\beta_0, \beta_l, \beta_k)$$

   to obtain the regression residual $\hat{\xi}_{it}(\beta_0, \beta_l, \beta_k)$ where we let $E[\omega_{it}|\omega_{i,t-1}] = g(\omega_{i,t-1}; \beta_\omega)$ for some parameter β_ω.[14]

2. Execute the CUE GMM using the moment condition

$$E\left[\hat{\xi}_{it}(\beta_0, \beta_l, \beta_k) \times \begin{pmatrix} 1 \\ l_{i,t-1} \\ l_{i,t-2} \\ k_{it} \\ k_{i,t-1} \end{pmatrix}\right] = 0.$$

We then conduct a Monte Carlo experiment using the same setup in Section 4. Figure 5 and Table 5 report the estimation results using this approach, conveying that this modified procedure is robust to different timing assumptions of labor input, and that the means of all three estimated coefficients, β_l, β_k and β_0, are very close to the true values.[15] Although estimating the constant together with β_l and β_k produces slightly larger standard deviations of the 1,000 estimates of β_l and β_k compared to the ones using the concentrated moments

[14] In practice, this means one can include higher order terms or approximating basis functions of $\hat{\omega}_{i,t-1}$ in the regression.

[15] In the estimation, we use different sets of initial values for β_l and β_k, as they are used in the previous experiments. Without loss of generality, we set the initial value of β_0 equal to zero.

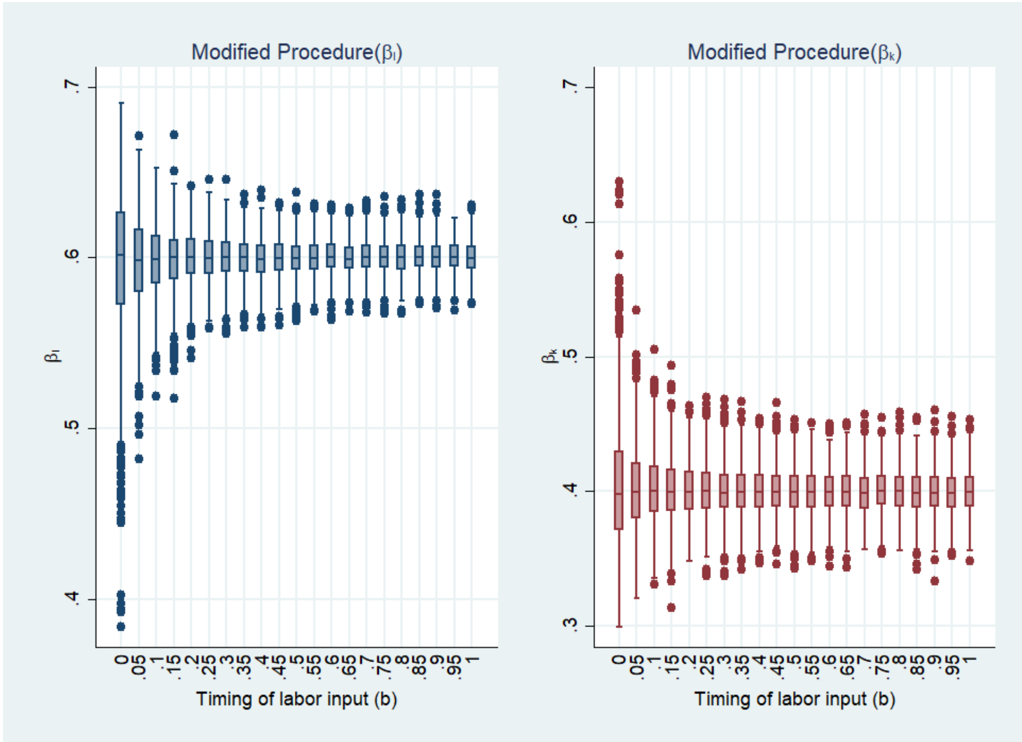

Figure 5: Modified Procedure with Augmented IV (Estimate β_0 jointly with (β_l, β_k))

(see Tables 3 and 4 for comparison) in the previous sections, the differences are rather small, implying that the estimated coefficients are all driven by the true minimum. Also, although the standard deviation of the estimates from this modified procedure increases marginally as b approaches zero, the procedure does not yield a spurious minimum, even at $b = 0$. This Monte Carlo exercise, therefore, illustrates that our proposed procedure can be generalized to estimate constant β_0 together with the other production function parameters.

5.4 Heterogeneous timing of labor input choice

Finally, we construct a DGP where the timing of labor input is heterogeneous across firms. Specifically, we randomly draw a value for each firm's timing of labor input b from a triangular distribution with support $[0, 1]$, and then generate the rest of data. All data settings other than this random timing are the same as in Section 4. We then simultaneously estimate production function parameters $(\beta_0, \beta_l, \beta_k)$ using the modified procedure in Section 5.3.

Figure 6 represents the estimation results when we vary the peak value of the triangular distribution from 0 to 1 with a stepsize of 0.05. For each distribution, we repeat the experiment 1,000 times to generate the distribution of the estimates. It is evident that our procedure is robust to heterogeneous timing of labor input choice. Also the performance of the proposed estimator is quite similar across different peak values of the triangular distribution, including

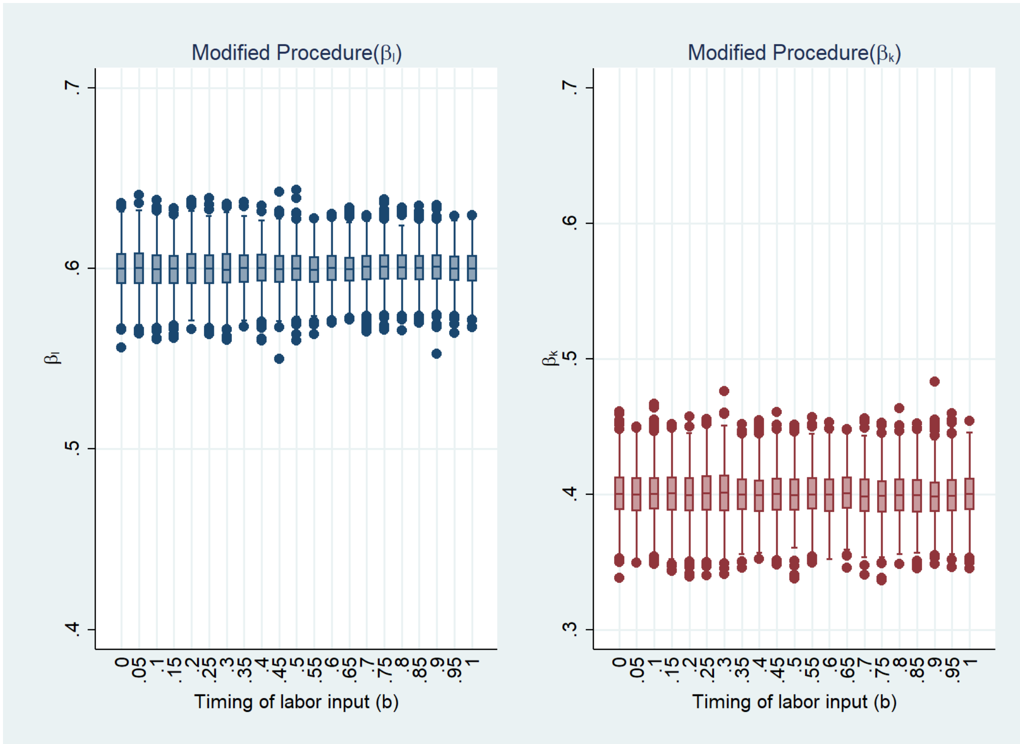

Figure 6: Heterogeneous Timing of Labor Choice: Modified Procedure with Augmented IV

Note: The box plots are generated using estimates from 1,000 simulated datasets for each timing of labor input. For each value of b, firms are assumed to choose their current labor input l_{it} at $t-b$, where b is drawn from a triangular distribution with support $[0, 1]$, and we vary the peak value of the distribution with a stepsize of 0.05. The true parameter values are $(\beta_l, \beta_k) = (0.6, 0.4)$.

values close to zero. This result suggests our proposed estimator will be useful in the more realistic data situation where the exact timing of input choices is unknown and may differ across firms.

6 Conclusion

The standard ACF approach allows flexible timing of labor input for when firms learn all or part of their productivity. However, we demonstrate that ACF's moment condition may suffer from weak identification, and the degree to which this causes issues varies by timing of labor input choice. We propose easy-to-implement modified procedures that remedy these issues, and provide Monte Carlo evidence. Since, in practice, the exact timing of input choices is unknown and may differ across firms, our proposed solution is valuable, as it fully incorporates the flexible framework but avoids the problem of weak identification.

References

[1] Ackerberg, D., K. Caves, and G. Frazer (2015), "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83, 2411-2451.

[2] Blundell, R., and S. Bond (2000), "GMM Estimation With Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19, 321-340.

[3] Gandhi, A., S. Navarro, and D. Rivers (2017), "On the Identification of Gross Output Production Functions," working paper.

[4] Hahn, J., G. Kuersteiner, and M. Mazzocco (2017), "Estimation with Aggregate Shocks," working paper.

[5] Hansen, P., J. Heaton, and A. Yaron (1996), "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economic Statistics*, 14, 262-280.

[6] Kim, K., Y. Luo, and Y. Su (2019), "A Robust Approach to Estimating Production Functions: Replication of the ACF procedure," *Forthcoming in Journal of Applied Econometrics*.

[7] Kim, K, A. Petrin, and S. Song (2016), "Estimating Production Functions with Control Functions When Capital Is Measured with Error," *Journal of Econometrics*, 190(2), 267-279.

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

[9] Marschak, J. and W. Andrews (1944), "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12, 143–205.

[10] 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.

[11] Olley, S. and A. Pakes (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263-1298.

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

Appendix

A Tables

		β_l		β_k	
Value of b	# of Replications	Mean	Std	Mean	Std
0.00	1,000	0.610	0.071	0.389	0.072
0.05	1,000	0.612	0.072	0.387	0.073
0.10	1,000	0.610	0.066	0.390	0.067
0.15	1,000	0.612	0.071	0.388	0.072
0.20	1,000	0.613	0.071	0.387	0.072
0.25	1,000	0.621	0.087	0.380	0.088
0.30	1,000	0.623	0.091	0.377	0.091
0.35	1,000	0.633	0.109	0.367	0.109
0.40	1,000	0.628	0.101	0.372	0.101
0.45	1,000	0.638	0.116	0.362	0.116
0.50	1,000	0.645	0.124	0.355	0.123
0.55	1,000	0.647	0.126	0.353	0.126
0.60	1,000	0.660	0.138	0.341	0.138
0.65	1,000	0.680	0.155	0.320	0.154
0.70	1,000	0.707	0.170	0.292	0.169
0.75	1,000	0.729	0.178	0.272	0.178
0.80	1,000	0.734	0.180	0.266	0.179
0.85	1,000	0.746	0.182	0.255	0.181
0.90	1,000	0.751	0.183	0.250	0.182
0.95	1,000	0.752	0.183	0.249	0.181
1.00	1,000	0.749	0.182	0.251	0.180

Table 1: ACF procedure with different timing of labor input

Note: The tables are generated using estimates from 1,000 simulated datasets for each timing of labor input. For each value of b, firms are assumed to choose their current labor input l_{it} at $t-b$. The true parameter values are $(\beta_l, \beta_k) = (0.6, 0.4)$.

		β_l		β_k	
Value of b	# of Replications	Mean	Std	Mean	Std
0.00	1,000	0.944	0.002	0.082	0.005
0.05	1,000	0.943	0.003	0.081	0.005
0.10	1,000	0.941	0.003	0.082	0.005
0.15	1,000	0.939	0.003	0.084	0.006
0.20	1,000	0.936	0.004	0.087	0.006
0.25	1,000	0.933	0.004	0.089	0.007
0.30	1,000	0.930	0.004	0.092	0.007
0.35	1,000	0.927	0.005	0.096	0.008
0.40	1,000	0.923	0.005	0.100	0.008
0.45	1,000	0.919	0.006	0.104	0.010
0.50	1,000	0.915	0.007	0.108	0.011
0.55	1,000	0.910	0.011	0.114	0.015
0.60	1,000	0.884	0.049	0.146	0.067
0.65	1,000	0.782	0.063	0.279	0.084
0.70	1,000	0.727	0.013	0.341	0.019
0.75	1,000	0.702	0.008	0.360	0.013
0.80	1,000	0.680	0.008	0.375	0.014
0.85	1,000	0.659	0.007	0.385	0.013
0.90	1,000	0.638	0.007	0.394	0.014
0.95	1,000	0.619	0.007	0.398	0.014
1.00	1,000	0.600	0.007	0.400	0.015

Table 2: ACF procedure with l_{it} as IV

		β_l		β_k	
Value of b	# of replications	Mean	Std	Mean	Std
0.00	1,000	0.599	0.016	0.400	0.007
0.05	1,000	0.598	0.015	0.400	0.006
0.10	1,000	0.599	0.013	0.400	0.006
0.15	1,000	0.599	0.012	0.400	0.006
0.20	1,000	0.600	0.012	0.400	0.006
0.25	1,000	0.600	0.011	0.400	0.006
0.30	1,000	0.600	0.010	0.400	0.006
0.35	1,000	0.599	0.010	0.401	0.006
0.40	1,000	0.599	0.010	0.401	0.006
0.45	1,000	0.600	0.009	0.400	0.006
0.50	1,000	0.600	0.009	0.400	0.006
0.55	1,000	0.600	0.009	0.400	0.006
0.60	1,000	0.600	0.009	0.400	0.005
0.65	1,000	0.600	0.009	0.400	0.006
0.70	1,000	0.600	0.008	0.400	0.005
0.75	1,000	0.600	0.008	0.400	0.006
0.80	1,000	0.600	0.009	0.400	0.006
0.85	1,000	0.600	0.008	0.400	0.006
0.90	1,000	0.600	0.008	0.400	0.006
0.95	1,000	0.600	0.008	0.400	0.006
1.00	1,000	0.600	0.009	0.400	0.006

Table 3: Modified Procedure with Augmented IV (Known β_0)

		β_l		β_k	
Value of b	# of replications	Mean	Std	Mean	Std
0.00	1,000	0.599	0.016	0.400	0.007
0.05	1,000	0.599	0.015	0.400	0.006
0.10	1,000	0.599	0.013	0.400	0.006
0.15	1,000	0.599	0.012	0.400	0.006
0.20	1,000	0.600	0.011	0.400	0.006
0.25	1,000	0.600	0.011	0.400	0.006
0.30	1,000	0.600	0.010	0.400	0.006
0.35	1,000	0.599	0.010	0.401	0.006
0.40	1,000	0.599	0.010	0.401	0.006
0.45	1,000	0.600	0.009	0.400	0.006
0.50	1,000	0.600	0.009	0.400	0.006
0.55	1,000	0.600	0.009	0.400	0.006
0.60	1,000	0.601	0.009	0.400	0.005
0.65	1,000	0.600	0.008	0.400	0.005
0.70	1,000	0.600	0.008	0.400	0.005
0.75	1,000	0.600	0.008	0.400	0.006
0.80	1,000	0.600	0.009	0.400	0.006
0.85	1,000	0.600	0.008	0.400	0.006
0.90	1,000	0.600	0.008	0.400	0.006
0.95	1,000	0.600	0.008	0.400	0.006
1.00	1,000	0.600	0.008	0.400	0.005

Table 4: Modified Procedure with Augmented IV (Estimating β_0 in the 1st stage)

		β_l		β_k		β_0	
Value of b	# of replications	Mean	Std	Mean	Std	Mean	Std
0.00	1,000	0.594	0.044	0.405	0.047	-0.007	0.062
0.05	1,000	0.597	0.027	0.402	0.031	-0.002	0.042
0.10	1,000	0.598	0.021	0.402	0.025	-0.002	0.036
0.15	1,000	0.599	0.018	0.401	0.023	-0.001	0.033
0.20	1,000	0.600	0.015	0.401	0.020	-0.001	0.029
0.25	1,000	0.600	0.014	0.401	0.020	-0.001	0.030
0.30	1,000	0.600	0.013	0.400	0.019	0.000	0.028
0.35	1,000	0.600	0.012	0.400	0.018	0.001	0.028
0.40	1,000	0.599	0.012	0.400	0.018	0.000	0.026
0.45	1,000	0.600	0.011	0.400	0.017	0.000	0.025
0.50	1,000	0.600	0.011	0.399	0.018	0.001	0.026
0.55	1,000	0.600	0.011	0.399	0.017	0.001	0.026
0.60	1,000	0.600	0.010	0.400	0.017	0.000	0.025
0.65	1,000	0.600	0.010	0.400	0.017	0.000	0.026
0.70	1,000	0.601	0.009	0.399	0.016	0.002	0.025
0.75	1,000	0.600	0.010	0.400	0.016	0.000	0.024
0.80	1,000	0.600	0.010	0.400	0.017	0.000	0.025
0.85	1,000	0.601	0.009	0.399	0.016	0.001	0.024
0.90	1,000	0.600	0.009	0.400	0.017	0.001	0.025
0.95	1,000	0.601	0.009	0.399	0.016	0.001	0.025
1.00	1,000	0.600	0.010	0.400	0.016	0.000	0.024

Table 5: Modified Procedure with Augmented IV (Estimating β_0 with β_l and β_k)

		β_l		β_k		β_0	
Value of b	# of replications	Mean	Std	Mean	Std	Mean	Std
0.00	1,000	0.600	0.012	0.400	0.018	-0.001	0.027
0.05	1,000	0.600	0.012	0.400	0.018	0.000	0.027
0.10	1,000	0.599	0.012	0.401	0.018	-0.001	0.027
0.15	1,000	0.599	0.012	0.400	0.019	0.000	0.028
0.20	1,000	0.600	0.011	0.400	0.017	0.001	0.026
0.25	1,000	0.599	0.012	0.401	0.019	-0.001	0.028
0.30	1,000	0.600	0.011	0.401	0.019	0.000	0.028
0.35	1,000	0.600	0.011	0.400	0.017	0.000	0.026
0.40	1,000	0.600	0.011	0.400	0.017	0.001	0.027
0.45	1,000	0.600	0.011	0.400	0.018	0.000	0.027
0.50	1,000	0.600	0.011	0.400	0.018	0.000	0.027
0.55	1,000	0.599	0.010	0.400	0.017	0.000	0.025
0.60	1,000	0.600	0.010	0.400	0.017	0.000	0.026
0.65	1,000	0.600	0.010	0.401	0.017	-0.001	0.025
0.70	1,000	0.600	0.010	0.399	0.017	0.002	0.026
0.75	1,000	0.601	0.010	0.399	0.017	0.002	0.026
0.80	1,000	0.600	0.010	0.400	0.017	0.000	0.025
0.85	1,000	0.600	0.011	0.399	0.018	0.002	0.027
0.90	1,000	0.601	0.010	0.398	0.017	0.003	0.025
0.95	1,000	0.600	0.010	0.399	0.017	0.001	0.024
1.00	1,000	0.600	0.010	0.400	0.017	-0.001	0.025

Table 6: Heterogeneous Timing of Labor Choice: Modified Procedure with Augmented IV

Note: This table is generated using estimates from 1,000 simulated datasets for each timing of labor input. For each value of b, firms are assumed to choose their current labor input l_{it} at $t-b$ where b is drawn from a triangular distribution with support $[0,1]$, and we vary the peak value of the distribution with a stepsize of 0.05. The true parameter values are $(\beta_0, \beta_l, \beta_k) = (0, 0.6, 0.4)$.