Production Function Estimation Robust to Flexible Timing of Labor Input

Kyoo il Kim
Michigan State University

Yao Luo
University of Toronto

Yingjun Su
Jinan University

May, 2019

Introduction

- Estimating production function better attributes the economic growth to its sources, and is important input to various research questions
- Fundamental challenge is due to the simultaneity of inputs and output decisions (Marschak and Andrews 1944)
- Least squares is inconsistent because labor and capital (or investment) depend on the unobserved productivity
- Literature to address simultaneity
  - Blundell and Bond (2000) use a dynamic panel method
  - Olley/Pakes (1996) and Levinsohn/Petrin (2003) use proxy variables: OP use investment and LP use intermediate input
  - Wooldridge (2009): a joint estimation approach of LP

  - Gandhi, Navarro, Rivers (2017) utilize a first order condition

- Proxy variable approaches like OP and LP make an implicit assumption on the timing of labor input
  - estimate labor and other static inputs coefficients in the first stage

- Ackerberg, Caves, Frazer (2015) show a functional dependence in the first stage of OP/LP

- ACF propose inverting the input demand being conditional on state variables including labor
  - "robust" to different timing assumptions on labor input
  - labor is allowed to be chosen at a point in time for which all or part of the productivity is learned by the firm

What We Do

- We show that ACF's moment condition is not robust to such timing assumptions
- ACF's moments suffer from either global identification failure or weak identification, yielding multiple solutions
  - degree of identification problem depends on timing of labor
- This failure of their robustness argument is a new finding, and is more serious than the "spurious minimum" problem
- First, we provide Monte Carlo evidences for the identification issues
- Second, we modify ACF procedure to produce robust methods
  - yield estimates robust to different timing assumptions of labor input

Plan of Talk

- Review production function estimation methods
- Discuss identification problems
- Monte Carlo experiments
  - ACF Procedure
  - Our Modified Procedure
- Discussion: Why our modifications help
- Conclude

Model and Estimation

- Briefly discuss production function estimation
- Value-added production function for a firm i at time t

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

  - l_{it} and k_{it} are labor and capital in log
  - ω_{it} is unobserved state variable (Total Factor Productivity) that impacts firm's decisions on inputs and production level
  - η_{it} is a pure i.i.d. shock
- OP/LP make an implicit assumption on the timing of labor input in their proxy variable approach

- LP assume labor is determined after intermediate input, so that intermediate input demand is given by

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

  - with scalar error and monotonicity, the inverse is $\omega_{it} = f_t^{-1}(k_{it}, m_{it})$, so m_t is a proxy for ω_{it}
  - Kim, Petrin, Song (2016), Collard-Wexler, De Loecker (2016), Hu, Huang, Sasaki (2017) to allow for multiple errors

- ACF argue this timing assumption can be restrictive, besides the functional dependence, since it does not allow l_{it} being a state variable

- To see this, suppose labor is given by $l_{it} = q_t(k_{it}, \omega_{it})$ and write

$$\begin{aligned} l_{it} &= q_t(k_{it}, f_t^{-1}(k_{it}, m_{it})) \\ &= q_t(k_{it}, g(f_{t-1}^{-1}(k_{i,t-1}, m_{i,t-1})) + \xi_{it}) \end{aligned}$$

  where $\omega_{it} = g(\omega_{i,t-1}) + \xi_{it}$ and ξ_{it} denotes an innovation

– given (k_{it}, m_{it}) or $(k_{it}, k_{i,t-1}, m_{i,t-1})$ there is no remaining exogenous variation of l_{it}

– functional dependence - multicollinearity

- To allow for more flexible assumption on labor input, ACF define intermediate input demand as

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

and, given the monotonicity, the inverse function

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

becomes the control in their setting

- This specification allows l_{it} to be determined before all or part of ω_{it} is realized at time t

ACF procedure

- Substituting into the production function,

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

- In the first stage, $\Phi_t(l_{it}, k_{it}, m_{it})$ is estimated by minimizing the sample counterpart of

$$E[(y_{it} - \Phi_t(l_{it}, k_{it}, m_{it}))^2]$$

- In the second stage, use the estimated control function Φ_t and assume a first order Markov process of the productivity

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

  to control for the endogeneity of inputs

- Because the composite shock $\xi_{it} + \eta_{it}$ is uncorrelated with firm's information at time $t-1$, obtain the moment condition

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

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

- Let $t-b$ denote a point in time the firm chooses labor input l_t for $b \in [0,1]$
- Depending on this timing assumption
  - use 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 $b \in [0,1)$
  - use $z_{it} = (1, k_{it}, l_{it}, l_{i,t-1}, \Phi_{t-1}(\cdot))$ if labor is chosen at $t-1$, i.e. labor depends on $\omega_{i,t-1}$, but not any part of ξ_{it}

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

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

- Let $w_{it} = \{(y_{it}, l_{it}) \cup z_{it}\}$ and write the moment condition

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

where $\theta = (\beta_0, \beta_l, \beta_k, \rho)$ and θ_0 denotes the true parameter

Concentrated ACF procedure

- ACF propose to concentrate out β_0 and ρ
- For a trial value of β_l and β_k, construct an estimate for $\beta_0 + \omega_{it}$ as

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

  where $\widehat{\Phi}_t(l_{it}, k_{it}, m_{it})$ is obtained from the first stage

- Then, run a regression of $\widehat{\beta_0 + \omega_{it}(\beta_l}, \beta_k)$ on its lag $\widehat{\beta_0 + \omega_{i,t-1}(}\beta_l, \beta_k)$ to obtain the residual as $\hat{\xi}_{it}(\beta_l, \beta_k)$
- Estimate β_l and β_k using the concentrated moments

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

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

- ACF also propose to use

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

if labor is chosen at time $t-1$

Spurious Minimum Problem?

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

Identification Problem

- Let $\omega_{i,t-1}$ evolve to $\omega_{i,t-b}$, at which point in time labor is chosen,

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

  and $Var(\xi_{it}^{B}) = \sigma_b^2$ such that $\xi_{it}^{B} = 0$ and $\sigma_b^2 = 0$ if $b = 0$

  - firms are allowed to choose l_{it} having less than perfect info about ω_{it} , and this info decreases as b increases
  - l_{it} has potential dynamic implications being chosen at period t, period $t-1$, or period $t-b$
  - ω_{it} affects intermediate input and production level

- We show the global identification problem happens only at a particular timing, in which l_{it} is determined only after all of ω_{it} is realized

- If labor is chosen at a point in time before ω_{it} is realized, the problem should not exist
  - However, our Monte Carlos find this failure still arises unless labor is determined quite ahead of the productivity
- On the other hand, if labor is chosen quite ahead of time, then ACF moments with $l_{i,t-1}$ suffer from weak identification
  - ACF may augment the current labor to IV
  - however, l_{it} is valid only if labor is determined before any part of ω_{it} is realized
- In practice we do not know when firms make labor input choices or different firms may have different timing of input choices, ACF should be used with some cautions

Setup

- The productivity is approximated as

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

- In 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 AR regression as

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

  and it becomes equivalent to $\widehat{\omega_{it}} = \rho\widehat{\omega_{i,t-1}} + \xi_{it}$ with $\alpha_0 = \beta_0(1 - \rho)$

- Note, however, including an intercept makes the residual $\xi_{it}(\beta_l, \beta_k)$ have mean zero by construction

- This is true whether or not the intercept is actually equal to the true $\alpha_0 = \beta_0(1 - \rho)$

Spurious Minimum: revisited

- To see how a spurious min arises, let $\tilde{\beta}_k = 0$ and $\tilde{\beta}_l = \beta_l + \beta_k = 1$

- Write $\Phi_{it} = \Phi_t(l_{it}, k_{it}, m_{it})$ for ease of notation and note

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

- Also, from the optimal labor input in the ACF's DGP, we can write $(1 - \beta_l) l_{it}$ as

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

  W_{it} denotes the wage and $\omega_{i,t-1}$ evolves to $\omega_{i,t-b}$, at which point in time firm chooses labor

- From $\omega_{it} = \rho^b \omega_{i,t-b} + \xi_{it}^B$, it follows that

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

- Two important implications for identification
  - First, the constant term $(-\ln\beta_l - \sigma^2/2)$ in the RHS is different from β_0
  - Second, the regression residual from the AR(1) regression at the spurious min using $\Phi_{it} - \tilde{\beta}_k k_{it} - \tilde{\beta}_l l_{it}$ is also different from ξ_{it}
- In the DGP the wage also follows an AR process as $\ln W_{it} = \rho_W \ln W_{i,t-1} + \xi_{it}^W$, then

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

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

  - innovation term ξ_{it}^W in the wage is independent of $(k_{it}, l_{i,t-1})$
  - as noted by ACF, the spurious parameter $(\tilde{\beta}_k = 0, \tilde{\beta}_l = 1)$ may solve the ACF's moment condition

- We, however, find this spurious minimum happens only at a particular timing of labor input
  - this is our nuanced argument, not previously noted by ACF
- Spurious residual has the composite error $\xi_{it}^B - \rho_W \xi_{i,t-1}^B + \xi_{it}^W$
- This does not satisfy the "spurious" moment condition with the IV $(k_{it}, l_{i,t-1})$ because $\xi_{i,t-1}^B$ (part of $\xi_{i,t-1}$, so part of $\omega_{i,t-1}$) is correlated with k_{it} unless $b = 0$
- On the other hand, $l_{i,t-1}$ does not violate the spurious moment for all $b \in [0, 1]$ implying $l_{i,t-1}$ is a weak IV

Spurious Minimum: revisited

- Spurious solution in the population arises only when $b = 0$ for which $\xi_{it}^B = \xi_{i,t-1}^B = 0$
- As long as $b > 0$, it should not arise because k_{it} is correlated with $\xi_{i,t-1}^B$
  - k_{it} deters the spurious moment condition from being satisfied
- In finite samples, the spurious min still can exist if the correlation between $\xi_{i,t-1}^B$ and k_{it} is small and/or b is close to zero
- Original ACF's DGP takes $b = 0.5$, so in this DGP there should be no spurious minimum
- However, even for this DGP the spurious minimum arises in finite samples (Kim, Luo, Su 2018)

Robustness Failure to different timing

- Although $l_{i,t-1}$ is a valid IV regardless of $b \in [0,1]$ but is potentially weak, yielding spurious solutions
  - $l_{i,t-1}$ is mean independent of the spurious residual for all $b \in [0,1]$
  - $k_{i,t}$ does not have this problem unless $b = 0$
- Look at this weak IV problem from the regression perspective
- Given a value of ρ, $l_{it} - \rho l_{i,t-1}$ can be regarded as the regressor for labor coefficient
- If the correlation between $l_{i,t-1}$ and the regressor $l_{it} - \rho l_{i,t-1}$ is not significant enough, then $l_{i,t-1}$ becomes an weak IV
  - differencing $l_{it} - \rho l_{i,t-1}$ may effectively dispense with the signal in the regressor and the noise dominates

- Similar argument from the panel data approach to production functions with a fixed effect (see e.g. Blundell and Bond 2000)
- From the labor input function,

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

  the only remaining exogenous variation of l_{it}, net of k_{it} and the lagged productivity, is the wage since they also appear in the regression equation

- Therefore, the differencing $l_{it} - \rho l_{i,t-1}$ at or near the spurious AR coefficient ρ_W will weaken the correlation between $l_{it} - \rho l_{i,t-1}$ and $l_{i,t-1}$, which may yield a weak IV problem

- As labor is determined at a point in time close to $t-1$, l_{it} becomes more relevant although it would violate the moment condition unless b is exactly equal to 1 because $E[\xi_{it} l_{it}] \neq 0$ as long as $b \neq 1$
- On the other hand, l_{it} is not subject to the spurious solution
  - l_{it} is always correlated with ξ_{it}^W, regardless of any b
- To summarize, $l_{i,t-1}$ is a valid IV regardless of $b \in [0,1]$ but is potentially weak, which may yield spurious solutions, while l_{it} is a strong IV but invalid unless $b = 1$
- This implies ACF suffers from the global identification failure if $b \to 0$, while it suffers from the weak identification if $b \to 1$
- ACF procedure with instrument l_{it} does not suffer from the global or weak identification, but it produces inconsistent estimates unless $b = 1$ for all firms

Monte Carlo Study

- Our Monte Carlos clearly illustrate these concerns
- We follow ACF's DGP while we vary the timing of labor such that firms choose their labor at $t - b$ for $b \in \{0, 0.05, 0.1, ..., 0.95, 1\}$
- Note that ACF's original DGP assumes $b = 0.5$
- Our design of the experiment is twofold
- First, we estimate the production function by replicating ACF's original procedure but varying the starting value, searching for the global min
- Second, after observing identification problems of the original ACF procedure, we propose modifications

- In all exercises, we try different starting values, i.e. (0, 1), (0.1, 0.9), (0.2, 0.8),...., (0.9, 0.1) for the search of the global min of the GMM objective function

- For each pair of initial values, we obtain an estimate of (β_l, β_k) that minimizes the objective function (local minimum)

- Compare the values of these local min and obtain the estimate that yields the global min for each simulated dataset

- Use Continuously Updated GMM estimator (CUE, Hansen, Heaton, Yaron 1996)

- Report the empirical distributions of estimates in box-plots by the timing of labor input

ACF Procedure with different timing

- Our first experiment to examine the finite sample performance of the original ACF under different timing of labor input
- For a given timing, we simulate 1000 datasets with 1000 firms for $T = 10$
- We estimate with the same ACF procedure
  - that is, we adopt the concentrated moment conditions with instruments (l_{t-1}, k_t) to estimate β_l and β_k
- We find the "spurious" minimum or weak identification clearly exists
- Degree to which the identification problem occurs depends on timing
  - $b < 0.7$, the spurious min arises more frequently
  - $b \geq 0.7$, the estimates are more dispersed, suggesting weak identification

- Also, S.D. of 1000 estimates get bigger as b increases, implying ACF moments become weak as labor is chosen at a point in time closer to $t-1$

- Identification problems prevail under all timing assumptions of labor choice

- The problem gets more severe as labor is determined at a point in time closer to $t-1$

- While using l_{it} as an IV does not produce a spurious min, it generates biased estimates unless $b=1$, i.e., when labor is chosen at time $t-1$

- This experiment confirms that while l_t is a relevant IV and is not subject to a spurious min, it is only valid when $b=1$

- From these experiments we conclude the spurious min or weak identification of the ACF procedure is not negligible in the finite samples

Our Modified Procedure

- We develop modified procedures that are robust to different timing of labor input as desired by ACF
- We exploit the fact that the total factor productivity evolves according to an AR(1) process without the intercept
- First we assume β_0 in the production function is known, helping us to gain some intuition about our modified procedures
- Then generalize our procedure to estimate this constant term

Procedure with Known β_0

1. We run the AR(1) regression of the productivity without the intercept

   (a) This is equivalent to the econometrician knowing $\beta_0 = 0$

2. We add further lagged input variables such as l_{t-2} and k_{t-1} to the instrument set

   (a) The further lagged input variables act as excluded instruments

   (b) Kim, Luo, Su (2018) discuss why the further lagged input variables are relevant instruments in the ACF setting

3. We also add "1" to the instrument set

   (a) This will ensure the regression residual $\hat{\xi}_{it}(\beta_l, \beta_k)$ to have the zero sample mean, which is the sample analogue of the innovation term ξ_{it} having the zero mean

Why "Constant" Matters

- In the ACF's DGP we have $\beta_0 = 0$ in the production function
- If run the AR(1) regression without the intercept, it helps the moment condition to yield a "true" solution deterring the spurious min
  - because the spurious min requires the AR(1) regression must have a non-zero intercept as $\tilde{\alpha} = (1 - \rho_W)(\ln \beta_l + \sigma^2/2)$ - an intercept different from $\alpha_0 = \beta_0(1 - \rho)$
- Without the intercept the "spurious" innovation term $\xi_{it}^B - \rho_W \xi_{it}^B + \xi_{it}^W$ in the "spurious" AR(1) regression would not have mean zero, which becomes inconsistent with the true DGP
- For our modified procedures, the AR(1) regression does not include the intercept (or removed)
  - estimate $\widehat{\omega}_{it} = \rho \widehat{\omega}_{i,t-1} + \xi_{it}$ to obtain the residual

- This regression without an intercept deters ξ_{it} from having zero mean at other parameter values
- Instead, we include "1" in the IV and ensure ξ_{it} has mean zero at the true parameter value
- If we included a constant in the regression, the residual would have zero mean for any parameter values
- Our augmented procedure is robust to different timing of labor
- S.D. of estimates for DGPs with higher b is even slightly smaller than DGPs with smaller b, while the original ACF was somewhat sensitive to different timing
- Adding the further lagged inputs in the IV clearly improves the estimates

Why Adding Instruments Helps

- The lagged capital $k_{i,t-1}$ is a relevant IV due to the construction of capital being accumulated through investment
- The intuition behind using $l_{i,t-2}$ as an additional IV is as follows
- In the ACF's DGP, one of the labor input determinants is the wage

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

- Because $\ln W_{it}$ follows an AR(1) process, further lagged labor inputs become relevant for l_{it}, unless the wage is included in the labor input function
- From our Monte Carlos, we find $l_{i,t-2}$ is a relevant IV, which helps for estimation, when it is augmented
- A puzzle arises because the ACF's labor input demand implies the further lagged labor should not be relevant

- To see this point, note that the ACF's labor input is given by $l_{it} = q_t(k_{it}, \omega_{it})$ and hence

$$\begin{aligned}
l_{it} &= q_t(k_{it}, f_t^{-1}(l_{it}, k_{it}, m_{it})) \\
&= q_t(k_{it}, g(f_{t-1}^{-1}(l_{i,t-1}, k_{i,t-1}, m_{i,t-1})) + \xi_{it}) \\
&= q_t(k_{it}, g(\Phi_{t-1}(l_{i,t-1}, k_{i,t-1}, m_{i,t-1}) - \beta_0 - \beta_l l_{i,t-1} - \beta_k k_{i,t-1}) + \xi_{it})
\end{aligned}$$

  where $m_{it} = f_t(l_{it}, k_{it}, \omega_{it})$ denotes intermediate input demand and $\omega_{it} = f_t^{-1}(l_{it}, k_{it}, m_{it})$ denotes its inverse

- Therefore, in the ACF setting, $l_{i,t-2}$ should not be relevant for the current labor given

$$(\Phi_{t-1}(l_{i,t-1}, k_{i,t-1}, m_{i,t-1}), k_{it}, k_{i,t-1}, l_{i,t-1})$$

- We, however, argue $l_{i,t-2}$ remains relevant for l_{it} in the ACF's DGP

- Because in the empirical inversion of the input demand (using $\widehat{\Phi}_t(l_{it}, k_{it}, m_{it})$)

the ACF's first stage does not include the wage and because the wage follows an AR process -wage is not data in many applications

- Usually in the production function estimation using a proxy variable we omit various other unobservable factors of the input demand, denoted by the "t" subscript
  - this may indicate market/industry structure, input prices (Olley and Pakes 1996), or other aggregate shocks (Hahn, Kuersteiner, Mazzocco 2017)
- In the ACF's setting this missing factor "t" is the wage
- Because of this "omitted" wage the lagged labor $l_{i,t-2}$ becomes a relevant IV for $l_{i,t}$ when the individual firm's wage is not included in the input demand function - due to data limitation

Estimating β_0 in the first step

- In practice, the constant β_0 in the production function is unknown
- Estimate sum of β_0 and mean of the productivity in the first stage, and remove it from $\widehat{\beta_0 + \omega_{it}}$
- We then obtain the de-meaned productivity $\widetilde{\omega}_{it}$
- Lastly, we use the regression $\widetilde{\omega}_{it} = \rho\widetilde{\omega}_{i,t-1} + \xi_{it}$ to obtain the innovation
- The remaining steps are following our modified procedure
- MC results show that our modified procedure is robust to different timing of labor input
  - Results are almost identical to those with known constant β_0
- The means of $\hat{\beta}_l$ and $\hat{\beta}_k$ are very close to the truth, and the S.D. of the estimates are much smaller than those from the original ACF

Estimating β_0 **jointly with** (β_l, β_k)

- General approach is to estimate $(\beta_0, \beta_l, \beta_k)$ simultaneously
- For this purpose we modify the concentrated moment condition

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

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

and run the AR(1) regression without a constant

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

and obtain the residual as $\hat{\xi}_{it}(\beta_0, \beta_l, \beta_k)$

2. Do the CUE GMM using the moment condition

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

- MC results show that this modified procedure is robust to different timing assumptions of labor as well as to heterogeneous timing of labor input by firms
- Means of all coefficients β_l, β_k and β_0 are very close to the truth
- Although estimating the constant with β_l and β_k generates slightly larger S.D. the differences are rather small

Conclusion

- Ackerberg, Caves, Frazer (2015) is not robust to different timing assumptions on labor input
- Monte Carlos shows ACF's moment condition suffers from either global identification failure or weak identification
- In practice we do not know when firms make their labor input or different firms may have different timing, our finding suggests that their procedure should be used with some cautions
- Our proposed method is robust to different timing, getting rid of both global and weak identification problems
- Future work: Estimation with "locally" invalid instruments