# Empirical IO

# Ackerberg et al. (2015), "Identification Properties of Recent Production Function Estimators"


Sergey Alexeev

University of Technology Sydney


March 23, 2017

# Big picture and Motivation

- Unobservables of production functions
- What they do vs what they should be doing
  - Can we reason the existence of a proxy?
  - ...and the time behavior of the unobserved heterogeneity?
- How to estimate all of it
  - First stage is to estimate the labor coefficient by regressing labor output and nonparametric function
  - Second stage estimate capital coefficient by using estimates from the first stage
- Identification assumption does not hold precluding consistency of the first stage
  - Labor coefficient can't be identified
  - Cf. the issue pointed by Dr. He with Robenson procedure (slide 31/38 Production Function Estimation)

# The Solution and the Plan

1. Conditional vs unconditional optimal input decisions
   - Review the assumptions
2. What unconditionality actually tells us? How restrictive possible DGPs?
   - Clarify the dependency
3. Can we do better?
   - Suggest weaker assumptions, use more information as proxy
4. How to apply it?
   - Construct estimation procedure
5. Compare the methods
   - Monte Carldo simulation

# Why model of dynamically optimizing firm is necessary?

- Panel data with fixed effects
  - Attenuation biases
- First order conditions
  - More flexible production function
  - No need for Hicks neutral shock only
  - Assumption of static FOCs
- IV
  - No need to "map" $k_{it}$ and $l_{it}$ to $p_{it}^{k}$ and $p_{it}^{l}$
  - Why do wages differ across firms at a point of time and within firms over time? (quality, slop, firms' skills)
- Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP)
  - Allow time varying unobservable
  - Allow for dynamic choices, yet without explicit solutions
  - No need for exogenous, across firm variation in input prices

# OP/LP typical assumptions

Environment assumptions

- A1: *Information set* $I_{it}$ includes $\{\omega_{i\tau}\}_{\tau=0}^{t}$, not $\{\omega_{i\tau}\}_{\tau=t+1}^{\infty}$, and $E[\varepsilon_{it} \mid I_{it}] = 0$
- A2: *First order Markov* $p(\omega_{it+1}|I_{it}) = p(\omega_{it+1}|\omega_{it})$, $p(\cdot)$ is known and stochastically increasing in $\omega_{it}$
- A3: *Timing of Input Choices* $k_{it} = \kappa(k_{it-1}, i_{it-1})$ and labor is non-dynamic. Note $k_{it} \in I_{it-1}$

Assumption on policy function

- A4: *Scalar Unobservable* $i_{it} = f_t(k_{it}, \omega_{it})$. Note an implicit assumption on heterogeneity of firms and differences across time
- A5: *Strict Monotonicity* $f_t(k_{it}, \omega_{it})$. Follows from A2

# Estimation prodecure: first stage

A4 and A5 imply $i_{it} = f_t(k_{it}, \omega_{it}) \rightarrow \omega_{it} = {f_t}^{-1}(k_{it}, i_{it})$

- NB! two arguments, no shocks

$$\begin{aligned} y_{it} &= \beta_o + \beta_k k_{it} + \beta_l l_{it} + {f_t}^{-1}(k_{it}, i_{it}) + \varepsilon_{it} \\ &= \beta_l l_{it} + \Phi_t(k_{it}, i_{it}) + \varepsilon_{it} \end{aligned} \quad (1)$$

- note that ${f_t}^{-1}$ is the solution to potentially complicated dynamic problem;
- keep in mind the definition of $\Phi_t(\cdot)$

$$E[\varepsilon_{it}|I_{it}] = E[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, i_{it})|I_{it}] = 0 \quad (2)$$

generates GMM $\hat{\beta}_l$ and $\hat{\Phi}_t(k_{it}, i_{it})$.

- If $\Phi_t$ is approximated independently then just OLS $y_{it}$ on $l_{it}$

# Estimation prodecure: second stage

A1 and A2 imply

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

where $E[\xi_{it}|I_{it-1}] = 0$. Then

$$\begin{aligned} y_{it} &= \beta_o + \beta_k k_{it} + \beta_l l_{it} + g(\omega_{it-1}) + \xi_{it} + \varepsilon_{it} \\ &= \beta_o + \beta_k k_{it} + \beta_l l_{it} ... \\ &+ g(\Phi_{t-1}(k_{it-1}, i_{it-1}) - \beta_o - \beta_k k_{it-1}) + \xi_{it} + \varepsilon_{it} \end{aligned}$$

with this moment condition procede as above

$$E[\xi_{it} + \varepsilon_{it}|I_{it-1}] = 0 \quad (3)$$

True by LIE $E[\xi_{it}|I_{it-1}] = 0$ and $E[\varepsilon_{it}|I_{it}] = 0 \Rightarrow E[\varepsilon_{it}|I_{it-1}] = 0$

# Intermediate input instead of investment decision

Consider $\beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} M_{it}^{\beta_m} e^{\omega_{it}} e^{\varepsilon_{it}}$
Then $y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \varepsilon_{it}$
By symmetry implies $m_{it} = f_t(k_{it}, \omega_{it})$ (A4b) that is strictly incresing in $\omega_{it}$ (A5b)
Note the advantages over A4 and A5

- Intermediat inputs are not-dynamic, proving existence does not require dynamic optimization
- Often $i_{it} = 0$, thus monotonicity often fails
- A4 rules firm-specific unobservables, such as capital adjustment costs and investment prices. While $m_{it}$ just like $l_{it}$ are non-dynamic and do not adjust across periods
    - A5 and serially correlated unobserved across firm heterogeneity

# Functional dependency

- Consider
  $\pi_{it} = p_y\{\beta_0 K_{it}^{\beta_k} L_{it}^{\beta_l} M_{it}^{\beta_m} e^{\omega_{it}} e^{\varepsilon_{it}}\} - p_m M_{it} - p_l L_{it} - r K_{it}$
- Then $p_y \frac{\partial f}{\partial M_{it}} = p_m \Rightarrow p_y \beta_o \beta_m K_{it}^{\beta_k} L_{it}^{\beta_l} M_{it}^{\beta_m - 1} e^{\omega_{it}} e^{\varepsilon_{it}} = p_m$
  - Note in log version $m_{it} = f_t(k_{it}, \omega_{it}, l_{it}; \{p\}, \{\beta\})$
- Take log $\beta_m + \beta_k K_{it} + \beta_l L_{it} + (\beta_m - 1) M_{it} + \omega_{it} + \varepsilon_{it} = \ln \frac{p_m}{p_y}$
- Invert for $\omega_{it}$ and combine with
  $y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \varepsilon_{it}$
- To get $y_{it} = \ln \frac{1}{\beta_m} + \ln \frac{p_m}{p_y} + m_{it} + \varepsilon_{it}$
- As "$R^2 \to 1$" $\beta_l$ (and $\beta_k$) disappear, they don't provide any new information. Everything is contained in $m_{it}, \{p\}, \{\beta\}$
  - moment condition is not informative on $\beta_l$ (Cf. (2) and (1))
- $E[\{l_{it} - E[l_{it}|k_{it}, m_{it}, t]\}\{l_{it} - E[l_{it}|k_{it}, m_{it}, t]\}']$ is p.d.

# Possible DGPs of $l_{it}$

Inspire youself with A4b and consider $l_{it} = h_t(k_{it}, \omega_{it})$,

- implies that $l_{it}$ as $m_{it}$ has no dynamic implications and chosed with full knowledge of $\omega_{it}$
- $l_{it} = g_t(k_{it}, f_t^{-1}(k_{it}, m_{it}))$; $b_l$ is inseparable from $\Phi_t(\cdot)$ in (1)

We need something like this $l_{it} = g_t(k_{it}, f_t^{-1}(k_{it}, m_{it}), \upsilon_{it})$

1. *i.i.d* optimization error in $l_{it}$ (not in $m_{it}$ (or $i_{it}$))
2. *i.i.d* shocks to the price of labor or output after $m_{it}$ (or $i_{it}$) is chosen but prior to $l_{it}$ being chosen
3. (in OP only) labor is non-dynamic and chosed at $t - b$ as a function of $\omega_{it-b}$, while $i_{it}$ is chosen at $t$

# 1 *i.i.d* optimization error

Consider $l_{it} = g_t(k_{it}, f_t^{-1}(k_{it}, m_{it}), \upsilon_{it})$; optimal level plus noise with $\upsilon_{it}$ *i.i.d.*

- Provides variation even after conditioning on $k_{it}, m_{it}$
- How about $m_{it}$? I.e. $l_{it} = g_t(k_{it}, f_t^{-1}(k_{it}, m_{it}, \eta_{it}), \upsilon_{it})$
  - nope, A4b is violated and $E[\upsilon_{it}\eta_{it}] \neq 0$
- Fine if planned material input are used; $\eta_{it} = 0$
- Won't work with unions, $\upsilon_{it}$ is not *i.i.d.* and "adds" into $m_{it}$
- Note that noise in observed labor that is independent from output (CME) is no good; attenuate $l_{it}$ and/or violate A4b
- Generalizes to OP

# 2 *i.i.d* shocks

Take $0 < b < 1$ and assume that $l_{it}$ is chosen at $t - b$ and $m_{it}$ at $t$

- $E[\varepsilon_{it}|I_{it}] = E[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, l_{it}, m_{it})|I_{it}] = 0$ no good
- $m_{it}$ generally depends on the previously optimally chosen $l_{it}$

How about the opposite; $m_{it}$ is chosen at $t - b$ and $l_{it}$ at $t$

- $E[\varepsilon_{it}|I_{it}] = E[y_{it} - \beta_l l_{it} - \Phi_t(k_{it}, m_{it})|I_{it}] = 0$ good
- $l_{it} = h_t(k_{it}, \omega_{it}, \wp_{it})$, where $\wp_{it}$ is unobservable *i.i.d*
- $\omega_{it}$ should be constant from $t - b$ to $t$; otherwise nonparametric function of $m_{it}$ and $k_{it}$ will not perfectly controle for $\omega_{it}$ in the moment condition.
- Generalizes to OP

# 3 imperfect knowledge if $w_{it}$

Keep "subperiod" $t-b$ in mind and assume $\omega$ "evolves through it"

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

$l_{it}$ is chosen at $t-b$, while $i_{it}$ is chosen at time $t$

- $l_{it} \in I_{it-b}$ and $i_{it} \in I_{it}$

Then $i_{it} = f_t(k_{it}, \omega_{it})$ and $l_{it} = g_t(\omega_{it-b}, k_{it})$

- Labor is chosen without perfect information about $\omega_{it}$, generating variation in $l_{it}$ conditional on $f_t^{-1}(k_{it}, i_{it})$
- Note that $l_{it}$ has to have $0$ dynamic implications
  - Otherwise, it would directly impact $i_{it}$

Can not be generalized to LP

# Alternative procedure: assumptions

Crunching numbers with OP or LP implies the faith in those DGPs

- Yet the assumptions can be relaxed

Consider "value-added" production function (ask Eamon)

- $y_{it}$ is proportional to $m_{it}$; $m_{t-b}$, $k_t$, $l_t$, what if $k_{t-b}$, $l_{t-b}$, $m_t$?

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

with the following assumptions

- A3c: $k_{it} = \kappa(k_{it-1}, i_{it-1})$,
  where $i_{it-1} \in I_{it-1}$, $l_{it} \in I_{it-1}$, or $(I_{it-b}, I_{it})$
  - $l_{it}$ affects current and future profit, e.g. hiring/firing costs
- A4c: $m_{it} = \tilde{f}_t(k_{it}, l_{it}, \omega_{it})$
  - More info to proxy $\omega$. Think $m_{it}$ is chosen after $l_{it}$
- A5c: $m_{it}$ is striclty increasing in $\omega_{it}$
  - $m_{it}$ is still non-dynamic

# Alternative procedure

First stage:

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

$$E[\varepsilon_{it}|I_{it}] = E[y_{it} - \widetilde{\Phi}_t(k_{it}, i_{it})|I_{it}] = 0$$

Second stage:

$$\begin{aligned} y_{it} &= \beta_o + \beta_k k_{it} + \beta_l l_{it} + \overbrace{g(\omega_{it-1}) + \xi_{it}}^{\omega_{it}} + \varepsilon_{it} \\ &= \beta_o + \beta_k k_{it} + \beta_l l_{it} ... \\ &+ g(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_o - \beta_k k_{it-1} - \beta_l l_{it-1}) + \xi_{it} + \varepsilon_{it} \end{aligned}$$

$$\begin{aligned} E[\xi_{it} + \varepsilon_{it}|I_{it-1}] = E[&y_{it} - \beta_o - \beta_k k_{it} - \beta_l l_{it} ... \\ &- g(\widetilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_o - \beta_k k_{it-1} - \beta_l k_{it-1})|I_{it-1}] ... \\ &= 0 \end{aligned}$$

# Estimation example

- $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$
- If $l_{it}$ is after $t-1$ then it correlates with $\xi_{it}$ and (3) fails
- 4 moment conditions (not 3 like in OP/LP); can actullay use 5
- and 4 parameters: $\beta_0$, $\beta_k$, $\beta_l$ and $\rho$

$$E\Bigg[(y_{it} - \beta_o - \beta_k k_{it} - \beta_l l_{it} -$$

$$\rho \cdot (\tilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_o - \beta_k k_{it-1} - \beta_l k_{it-1}))$$

$$\otimes \left( \begin{array}{c} 1 \\ k_{it} \\ l_{it-1} \\ \tilde{\Phi}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) \end{array} \right) \Bigg] = 0$$

# Discussion and extensions

- Why suggested procedure is so great
  - No need to be believe in one on those three DGPs
  - Labor is dynamic
  - Serially correlated wage conditions (A4c generalizes A4b)
- Investment function approach
- Joint estimation
- Relation to dynamic panel methods

# Monte Carlo results

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

[a] 1000 replications. True values of $\beta_l$ and $\beta_k$ are 0.6 and 0.4, respectively. Standard deviations reported are of parameter estimates across the 1000 replications.

# References


Ackerberg, D. A. et al. (2015). "Identification Properties of Recent Production Function Estimators". In: *Econometrica* 83.6, pp. 2411–2451.

Levinsohn, J. and A. Petrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables". In: *The Review of Economic Studies* 70.2, pp. 317–341. ISSN: 00346527, 1467937X. URL: http://www.jstor.org/stable/3648636.

Olley, G. S. and A. Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry". In: *Econometrica* 64.6, pp. 1263–1297. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2171831.