ECO 310: Empirical Industrial Organization

Lecture 4: Production Functions: Estimation Methods

Victor Aguirregabiria (University of Toronto)

October 1st, 2018

Outline: Estimation Methods

1. Input prices as IVs
2. Fixed Effects estimator
3. Fixed Effects - Cochrane–Orcutt estimator
4. Arellano-Bond estimator
5. Panel Data: System estimator
6. Control Function: Olley-Pakes estimator
7. Control Function: Levinsohn-Petrin estimator
8. Using First Order Conditions

1. Input Prices as Instruments

Input prices as IVs

- If input prices r_{it} are observable (wages, cost of capital, fuel and energy prices), then under the assumption that they are not correlated with TFP, $E\left(\omega_{it}\ r_{it}\right) = 0$, we can use them as instruments.
- This approach has several **limitations/problems**.
- **Problem (1)**. Firms in the same industry typically use very similar type of inputs (labor, capital equipment, energy, materials) and they buy these inputs in the same input markets. If these input markets are competitive, the input prices are the same for all the firms in the industry:

$$r_{it} = r_t \quad \text{for every firm } i$$

- If input prices vary only over time, they are perfectly collinear with time-dummies in the PF. No valid instruments.

Input prices as IVs (2)

- **Problem (2)**. When input prices have cross-sectional variation, it could be because endogenous reasons.
- (a) Inputs markets are not competitive and firms with higher productivity pay higher prices. Then, $cov\left(\omega_{it}, r_{it}\right) \neq 0$, making input prices not a valid instrument.
- (b) Firms may be using different types of labor or capital inputs, with different qualities. This difference in the quality of inputs is part of the log-TFP. Then, $cov\left(\omega_{it}, r_{it}\right) \neq 0$, making input prices not a valid instrument.

Input prices as IVs (3)

- An ideal situation for using input prices as IVs is when firms in the same industry product in different geographic markets where the input markets are competitive.
- The variation in input prices over gegraphic market is due to different conditions on the supply of inputs (e.g., labor supply, better access to materials) and not to differences in productivity.

2. Fixed Effects estimator

Fixed-Effects (FE) estimator

- Consider the PF:

$$y_{it} = \alpha_L \ \ell_{it} + \alpha_K \ k_{it} + \omega_{it} \tag{1}$$

- Let's first **define the FE (or Within-Groups) estimator** and then we will show under which conditions this estimator provides unbiased (consistent) estimates of parameters α_L and α_K.

- If, for each firm i, we average equation (1) considering all the years observations, we have the equation:

$$\bar{y}_i = \alpha_L \ \overline{\ell}_i + \alpha_K \ \overline{k}_i + \overline{\omega}_i \tag{2}$$

where:

$$\bar{y}_i = \frac{\sum_{t=1}^{T} y_{it}}{T}; \ \overline{\ell}_i = \frac{\sum_{t=1}^{T} \ell_{it}}{T}; \ \overline{k}_i = \frac{\sum_{t=1}^{T} k_{it}}{T}; \ \overline{\omega}_i = \frac{\sum_{t=1}^{T} \omega_{it}}{T}$$

FE estimator (2)

- If we subtract equation (2) to equation (1), we have:

$$(y_{it} - \bar{y}_i) = \alpha_L \left(\ell_{it} - \overline{\ell}_i\right) + \alpha_K \left(k_{it} - \bar{k}_i\right) + (\omega_{it} - \overline{\omega}_i) \qquad (3)$$

- This equation is named the **Fixed-Effects** (or the Within-Groups) **transformation of the model**.
- The **FE estimator** is OLS applied to FE transformed model.
- For instance, if we had only one input, say labor, the FE estimator of α_L would be:

$$\widehat{\alpha_L^{OLS}} = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it} - \overline{y}_i) \left(\ell_{it} - \overline{\ell}_i\right)}{\sum_{i=1}^{N} \sum_{t=1}^{T} \left(\ell_{it} - \overline{\ell}_i\right)^2}$$

FE estimator in Stata: Implementation 1

- Let `logy`, `logn`, `logk`, `id`, `year`, be the variables.
- First, we construct the within-group transformation of the variables:

```
egen mlogy = mean(logy), by(id)
gen wlogy = logy - mlogy

egen mlogn = mean(logn), by(id)
gen wlogn = logn - mlogn

egen mlogk = mean(logk), by(id)
gen wlogn = logn - mlogn
```

Implementation 1

Then, run OLS using within-groups transformed variables.

```
reg wlogy wlogn wlogk
```

-

Source	SS	df	MS
Model	211.639421	2	105.81971
Residual	81.9304197	4,069	.020135271
Total	293.56984	4,071	.072112464

Number of obs	=	4,072
F(2, 4069)	=	5255.44
Prob > F	=	0.0000
R-squared	=	0.7209
Adj R-squared	=	0.7208
Root MSE	=	.1419

wlogy	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
wlogn	.5995702	.012516	47.90	0.000	.575032	.6241084
wlogk	.3190445	.0092708	34.41	0.000	.3008687	.3372203
_cons	4.53e-09	.0022237	0.00	1.000	-.0043597	.0043597

FE estimator: Implementation 2

- Define *i* and *t* in panel data. Then, we apply FE estimator:

```
xtset id year
xtreg logy logn logk, fe
```

```
Fixed-effects (within) regression               Number of obs      =     4,072
Group variable: id                              Number of groups   =       509

R-sq:                                           Obs per group:
     within  = 0.7209                                         min =          8
     between = 0.9729                                         avg =        8.0
     overall = 0.9683                                         max =          8

                                                F(2,3561)          =   4599.32
corr(u_i, Xb)  = 0.4651                         Prob > F           =    0.0000

------------------------------------------------------------------------------
        logy |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        logn |   .5995702    .013379    44.81   0.000     .5733389    .6258015
        logk |   .3190445     .00991    32.19   0.000     .2996146    .3384744
       _cons |   3.512405   .0404574    86.82   0.000     3.433083    3.591727
-------------+----------------------------------------------------------------
     sigma_u |  .37142884
     sigma_e |  .15168289
         rho |  .85706581   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(508, 3561) = 36.48                  Prob > F = 0.0000
```

FE estimator: Implementation 2

- We can also include year dummies:

```
xtreg logy logn logk i.year, fe
```

```
Fixed-effects (within) regression               Number of obs      =     4,072
Group variable: id                              Number of groups   =       509

R-sq:                                           Obs per group:
     within  = 0.7379                                         min =          8
     between = 0.9706                                         avg =        8.0
     overall = 0.9661                                         max =          8

                                                F(9,3554)          =   1111.47
corr(u_i, Xb)  = 0.5988                         Prob > F           =    0.0000
```

logy	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logn	.6544609	.0144048	45.43	0.000	.6262184	.6827034
logk	.2329072	.013637	17.08	0.000	.2061702	.2596443
year						
1983	-.0376406	.0093042	-4.05	0.000	-.0558828	-.0193985
1984	-.0076445	.0096071	-0.80	0.426	-.0264805	.0111914
1985	-.0234513	.0100955	-2.32	0.020	-.0432449	-.0036578
1986	-.0136103	.0105543	-1.29	0.197	-.0343034	.0070829
1987	.0314121	.0108748	2.89	0.004	.0100907	.0527335
1988	.0753576	.0111072	6.78	0.000	.0535805	.0971347
1989	.0764164	.0118166	6.47	0.000	.0532485	.0995844

Consistency of FE estimator

- The FE estimator is OLS in the regression equation

$$(y_{it} - \bar{y}_i) = \alpha_L \left(\ell_{it} - \overline{\ell}_i\right) + \alpha_K \left(k_{it} - \bar{k}_i\right) + (\omega_{it} - \overline{\omega}_i) \tag{3}$$

- As any OLS estimator, it is consistent if the error term is not correlated with the regressors. In this case, this implies:

$$\mathbb{E}\left[(\omega_{it} - \overline{\omega}_i) \left(\ell_{it} - \overline{\ell}_i\right)\right] = \mathbb{E}\left[(\omega_{it} - \overline{\omega}_i) \left(k_{it} - \overline{k}_i\right)\right] = 0$$

- We now present two **assumptions** on the unobserved log-TFP that imply consistency of the FE estimator (with time dummies).

Consistency of FE estimator (2)

- **Assumption FE-1.** Log-TFP has the following structure:

$$\omega_{it} = \eta_i + \delta_t + u_{it}$$

- η_i is interpreted as managerial ability, or a different technology that is constant over time.
- δ_t represents productivity that affect in the same way to all the firms in the industry.
- u_{it} is a firm-specific transitory shock.

Consistency of FE estimator (3)

- **Assumption FE-2.** The firm-specific transitory shock, u_{it}, is not correlated over time and it is realized after the firm chooses the amount of inputs at period t.
- u_{it} is a "surprised" that is realized after the firm has chosen inputs. For any two periods t and s, u_{it} is not correlated with inputs ℓ_{is} and k_{is}.

Consistency of FE estimator (4)

- Under assumptions FE-1 we have that:

$$\overline{\omega}_i = \eta_i + \overline{\delta} + \overline{u}_i$$

such that:

$$\omega_{it} - \overline{\omega}_i = \delta_t - \overline{\delta} + u_{it} - \overline{u}_i$$

- If we control for $\delta_t - \overline{\delta}$ using time dummies, the remaining error term is $u_{it} - \overline{u}_i$.

- Under Assumption FE-2, the error term $u_{it} - \overline{u}_i$ is not correlated with the regressors $\left(\ell_{it} - \overline{\ell}_i\right)$ and $(k_{it} - \overline{k}_i)$ because, for any two periods t and s, u_{it} is not correlated with inputs ℓ_{is} and k_{is}.

- Under FE-1 and FE-2, the FE estimator is unbiased / consistent.

3. Cochrane–Orcutt estimator

Cochrane–Orcutt estimator

- The assumption that the firm-specific transitory shock is not serially correlated (and fully unknown to the firm at period t) is quite strong.
- This assumption is testable (Arellano-Bond test for serial correlation). If rejected, this assumption can be relaxed.
- Suppose that we maintain assumption FE-1 but we replace assumption FE-2 with the following.
- **Assumption FE-CO.** The firm-specific transitory shock, u_{it}, follows an Autorregressive-1 process, AR(1):

$$u_{it} = \rho \ u_{i,t-1} + a_{it}$$

where ρ is a parameter, and a_{it} is not correlated over time and it is realized after the firm chooses the amount of inputs at period t.

Cochrane–Orcutt estimator (2)

- In this model where u_{it} is serially correlated, the standard FE estimator is inconsistent (biased) because $u_{it} - \overline{u}_i$ is correlated with the regressors.
- However, we can define a new version of the FE estimator (Cochrane-Orcutt FE) that is consistent under these conditions and the additional condition that the number of periods T is large.

Cochrane–Orcutt estimator (3)

- Consider the PF at periods t and $t-1$ under assumption FE-1:

$$\begin{array}{rcl} y_{it} & = & \alpha_L \; \ell_{it} + \alpha_K \; k_{it} + \eta_i + \delta_t + u_{it} \\ y_{it-1} & = & \alpha_L \; \ell_{it-1} + \alpha_K \; k_{it-1} + \eta_i + \delta_{t-1} + u_{it-1} \end{array}$$

- Multiplying the equation at $t-1$ by ρ and subtracting it to the equation at period t, we get:

$$\begin{array}{rcl} y_{it} - \rho \; y_{it-1} & = & \alpha_L \left[\ell_{it} - \rho \; \ell_{it-1} \right] + \alpha_K \left[k_{it} - \rho \; k_{it-1} \right] \\ & + & \left[1 - \rho \right] \eta_i + \left[\delta_t - \rho \; \delta_{t-1} \right] + a_{it} \end{array}$$

  because $u_{it} - \rho \; u_{it-1} = a_{it.}$

- This is called a **quasi-difference transformation**.

Cochrane–Orcutt estimator (4)

- The quasi-difference transformation can be written as:

$$\begin{aligned} y_{it} &= \beta_1 \; y_{i,t-1} + \beta_2 \; \ell_{it} + \beta_3 \; \ell_{it-1} + \beta_4 \; k_{it} + \beta_5 \; k_{it-1} \\ &+ \eta_i^* + \delta_t^* + a_{it} \end{aligned} \tag{4}$$

  with $\beta_1 = \rho$, $\beta_2 = \alpha_L$, $\beta_3 = -\rho\alpha_L$, $\beta_4 = \alpha_K$, $\beta_5 = -\rho\alpha_K$, and $\eta_i^* = [1 - \rho]\, \eta_i$, and $\delta_t^* = \delta_t - \rho\; \delta_{t-1}$.

- Note that given the β parameters we can obtain the parameters ρ, α_L, and α_K. In fact, there are additional (over-identifying restrictions):

$$-\beta_3/\beta_2 = -\beta_5/\beta_4 = \beta_1$$

- Now, under assumption FE-CO, in equation (2), the transitory shock a_{it} is not correlated with the inputs.

Cochrane–Orcutt estimator (5)

- Consider equation (4) in firm-specific means:

$$
\begin{array}{rcl}
\overline{y}_i & = & \beta_1 \, \overline{y}_{i(-1)} + \beta_2 \, \overline{\ell}_i + \beta_3 \, \overline{\ell}_{i(-1)} + \beta_4 \, \overline{k}_i + \beta_5 \, \overline{k}_{i(-1)} \\
& + & \eta_i^* + \overline{\delta}^* + \overline{a}_i
\end{array}
$$

- And in deviations with respect to firm-specific means:

$$
\begin{array}{rcl}
y_{it} - \overline{y}_i & = & \beta_1 \, \left[y_{it-1} - \overline{y}_{i(-1)} \right] + \beta_2 \left[\ell_{it} - \overline{\ell}_i \right] + \beta_3 \, \left[\ell_{it-1} - \overline{\ell}_{i(-1)} \right] \\
& + & \beta_4 \left[k_{it} - \overline{k}_i \right] + \beta_5 \, \left[k_{it-1} - \overline{k}_{i(-1)} \right] \\
& + & (\delta_t^* - \overline{\delta}^*) + (a_{it} - \overline{a}_i)
\end{array}
$$

- The FE-Cocharne-Orcutt estimator is applying OLS to this equation.

Cochrane–Orcutt estimator: Implementation

```
xtreg logy logn logk l.logy l.logn l.logk i.year, fe
```

```
Fixed-effects (within) regression               Number of obs     =      3,563
Group variable: id                              Number of groups  =        509

R-sq:                                           Obs per group:
     within  = 0.7825                                         min =          7
     between = 0.9879                                         avg =        7.0
     overall = 0.9847                                         max =          7

                                                F(11,3043)        =     995.10
corr(u_i, Xb)  = 0.7191                         Prob > F          =     0.0000
```

logy	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
logn	.4880013	.0166747	29.27	0.000	.4553065	.5206961
logk	.1765454	.0178288	9.90	0.000	.1415877	.2115032
logy						
L1.	.4039344	.015273	26.45	0.000	.3739879	.4338808
logn						
L1.	-.0231194	.0192464	-1.20	0.230	-.0608566	.0146179
logk						
L1.	-.1305487	.0164086	-7.96	0.000	-.1627218	-.0983757

Cochrane–Orcutt estimator: Large T condition

- **IMPORTANT NOTE:** The FE-Cochrane-Orcutt is consistent (asymp. unbiased) only when T is large, e.g., larger than 30 or 40 periods.
- Note that under condition FE-OC, we have that $(a_{it} - \overline{a}_i)$ is not correlated with regressors $\left[\ell_{it} - \overline{\ell}_i\right]$, $\left[\ell_{it-1} - \overline{\ell}_{i(-1)}\right]$, $\left[k_{it} - \overline{k}_i\right]$, and $\left[k_{it-1} - \overline{k}_{i(-1)}\right]$.
- However, even under this condition, we have that $(a_{it} - \overline{a}_i)$ is correlated with regressor $\left(y_{it-1} - \overline{y}_{i(-1)}\right)$.
- Note that y_{it-1} depends on $a_{i,t-1}$ and that $a_{i,t-1}$ is part of $\overline{a}_i$.
- This correlation goes to zero as T becomes large.

4. Arellano-Bond estimator

Arellano-Bond estimator

- Assumption FE-2 (or for that matter FE-CO) has two parts:
  - FE-2(a): u_{it} is not serially correlated.
  - FE-2(b): u_{it} is not known to the firm when it decides the amounts of inputs.
- In most applications, the stronger of the two assumptions is FE-2(b).
- We now present a panel data estimator that relaxes assumption FE-2(b).

Arellano-Bond estimator (2)

- We maintain assumptions FE-1, $\omega_{it} = \eta_i + \delta_t + u_{it}$, and FE-2(a), u_{it} is not serially correlated.
- Define the variables in first differences: $\Delta y_{it} = y_{it} - y_{it-1}$; $\Delta \ell_{it} = \ell_{it} - \ell_{it-1}$; etc.
- And consider the PF in first differences (equation at period t minus equation at period $t-1$):

$$\Delta y_{it} = \alpha_L \ \Delta \ell_{it} + \alpha_K \ \Delta k_{it} + \Delta \delta_t + \Delta u_{it}$$

- We have removed the term η_i from the error term, and we can control for the term $\Delta \delta_t$ by including time-dummies.
- But we still have the term Δu_{it} that is correlated with the regressors $\Delta \ell_{it}$ and Δk_{it}.

Arellano-Bond estimator (3)

$$\Delta y_{it} = \alpha_L \ \Delta \ell_{it} + \alpha_K \ \Delta k_{it} + \Delta \delta_t + \Delta u_{it}$$

- Consider the following general models for demand of capital and labor inputs:

$$(\text{LD}) \quad \ell_{it} = f_L(\ell_{i,t-1}, k_{i,t-1}, \omega_{it}, r_{it})$$

$$(\text{KD}) \quad k_{it} = f_K(\ell_{i,t-1}, k_{i,t-1}, \omega_{it}, r_{it})$$

- This means that ℓ_{it} and k_{it} depends on the current and the past histories of the transitory shocks: u_{it}, u_{it-1}, $u_{i,t-2}$, ...;
- But not on future shocks: u_{it+1}, u_{it+1}, ...

Arellano-Bond estimator (4)

$$\Delta y_{it} = \alpha_L \ \Delta \ell_{it} + \alpha_K \ \Delta k_{it} + \Delta \delta_t + \Delta u_{it}$$

- This implies that $\ell_{i,t-2}$ and $k_{i,t-2}$ are valid instruments in this equation.
- They are **Relevant**: $\Delta \ell_{it}$ and Δk_{it} are correlated with $\ell_{i,t-2}$ and $k_{i,t-2}$.
- They are **not correlated with the error term**:

$$\mathbb{E}\left[\ell_{it-2} \ \Delta u_{it}\right] = \mathbb{E}\left[\ell_{it-2} \ u_{it}\right] - \mathbb{E}\left[\ell_{it-2} \ u_{it-1}\right] = 0$$

Arellano-Bond estimator (5)

- This idea implies many moment restrictions that can be used to estimate α_L, α_K, and $\Delta\delta_t$:

$$E\left(\ell_{i,t-j}\ \Delta u_{it}\right) = 0 \quad \text{for } t = 3, ..., T; \ \text{ and } j \leq t-2$$

$$E\left(k_{i,t-j}\ \Delta u_{it}\right) = 0 \quad \text{for } t = 3, ..., T; \ \text{ and } j \leq t-2$$

$$E\left(y_{i,t-j}\ \Delta u_{it}\right) = 0 \quad \text{for } t = 3, ..., T; \ \text{ and } j \leq t-2$$

- The Arellano-Bond estimator exploits all these restrictions optimally: optimal weighting; optimal Generalized Method of Moments (GMM) estimator.

Arellano-Bond estimator: Implementation

- The command `xtabond2` in Stata implements the Arellano-Bond estimator.
- It can be applied also to the model where u_{it} follows an autorregressive process.

5. System GMM estimator

System GMM

- When labor and capital inputs are strongly correlated, the Arellano-Bond estimator suffers of a **weak instruments problem**: low correlation between instruments and endogenous variables, and imprecise estimates.
- Note that if ℓ_{it} and k_{it} follow "random walks" then $\Delta\ell_{it}$ and Δk_{it} are not serially correlated and therefore they are not correlated with the instruments ℓ_{it-2} and k_{it-2}.
- For these cases, Blundell-Bond derive additional restrictions that help to identify the PF.

Blundell and Bond (2001)

- They show that if the model is stationary, then $\Delta \ell_{it}$ and Δk_{it} are not correlated with η_i.
- Therefore, in the PF in levels:

$$y_{it} = \alpha_L \ \ell_{it} + \alpha_K \ k_{it} + \delta_t + \eta_i + u_{it}$$

  we have that $\Delta \ell_{it-1}$ and Δk_{it-1} are not correlated with error term $\eta_i + u_{it}$.
- This implies that $\Delta \ell_{it-1}$ and Δk_{it-1} are valid instruments in the equation in levels.
- **Relevant**. ℓ_{it} and k_{it} are correlated with $\Delta \ell_{it-1}$ and Δk_{it-1} (even when inputs follow random walks).
- **No correlation with error:** $E\left(\Delta \ell_{it-1}\left[\eta_i + u_{it}\right]\right) = E\left(\Delta k_{it-1}\left[\eta_i + u_{it}\right]\right) = 0$

Blundell and Bond (2001)

- The Blundell-Bond **System GMM estimator** combines these moment restrictions and Arellano-Bond moment restrictions in an optimal way to obtain an efficient estimator.

Blundell and Bond (2001): Results

509 manufacturing firms; 1982-89				
Parameter	**OLS-Levels**	**WG**	**AB-GMM**	**SYS-GMM**
α_L	0.538	0.488	0.515	0.479
	(0.025)	(0.030)	(0.099)	(0.098)
α_K	0.266	0.199	0.225	0.492
	(0.032)	(0.033)	(0.126)	(0.074)
ρ	0.964	0.512	0.448	0.565
	(0.006)	(0.022)	(0.073)	(0.078)
Sargan (p-value)	-	-	0.073	0.032
m2	-	-	-0.69	-0.35
Constant RS (p-v)	0.000	0.000	0.006	0.641

7. Control Function: Olley and Pakes estimator

Control Function Methods

- Olley & Pakes (1996; OP) and Levinsohn & Petrin (2003; LP) are **control function methods**.
- Instead of looking for instruments for K and L, we look for observable variables that can "control for" (or proxy) unobserved TFP.
- The control variables should come from a model of firm behavior.
- Note: Both OP and LP assume that labor is perfectly flexible input. This assumption is completely innocuous for their results. To emphasize this point, I present here versions of OP and LP that treat labor as a potentially dynamic input.

Olley and Pakes (OP)

- Consider the following model of simultaneous equations:

$$\boxed{\begin{array}{ll} (PF) & y_{it} = \alpha_L \ \ell_{it} + \alpha_K \ k_{it} + \omega_{it} + e_{it} \\ \\ (LD) & \ell_{it} = f_L\left(\ell_{i,t-1}, k_{it}, \omega_{it}, r_{it}\right) \\ \\ (ID) & i_{it} = f_K\left(\ell_{i,t-1}, k_{it}, \omega_{it}, r_{it}\right) \end{array}}$$

  (LD) & (ID): firms' optimal labor and investment given state variables $(\ell_{i,t-1}, k_{it}, \omega_{it}, r_{it})$; r_{it} = input prices.

- OP consider the following assumptions:

$$\begin{array}{ll} (OP-1) & f_K\left(\ell_{i,t-1}, k_{it}, \omega_{it}, r_{it}\right) \text{ is invertible in } \omega_{it} \\ (OP-2) & \text{No cross-sectional variation in } r_{it}\text{: } r_{it} = r_t. \\ (OP-3) & \omega_{it} \text{ follows a first order Markov process.} \\ (OP-4) & k_{it} \text{ is decided at } t-1\text{: } k_{it} = (1-\delta) k_{i,t-1} + i_{i,t-1} \end{array}$$

Olley and Pakes (2)

- OP method deals both with the simultaneity problem and with the selection problem due to endogenous exit.
- It doesn't deal with potential measurement error in inputs.
- OP method proceeds in **two stages**.
- **First stage:** estimates α_L [Assumptions (OP-1) and (OP-2) are key]; and the **second stage** estimates α_K [Assumptions (OP-3) and (OP-4) are key].

Olley and Pakes First Stage

- Assumptions (OP-1) and (OP-2) imply that the investment equation is invertible in ω_{it}:

$$\omega_{it} = f_K^{-1}(\ell_{i,t-1}, k_{it}, i_{it}, r_t)$$

- Solving this equation in the PF we have:

$$\begin{aligned} y_{it} &= \alpha_L\ \ell_{it} + \alpha_K\ k_{it} + f_K^{-1}(\ell_{i,t-1}, k_{it}, i_{it}, r_t) + e_{it} \\ &= \alpha_L\ \ell_{it} + \phi_t(\ell_{i,t-1}, k_{it}, i_{it}) + e_{it} \end{aligned}$$

- This is a **partially linear model**. Parameter α_L and functions $\phi_1(.)$, ..., $\phi_T(.)$ can be estimated using **semiparametric methods**.
- A possible method is Robinson's method (1988). OP use an $n-th$ order polynomial to approximate the ϕ_t functions.

Olley and Pakes First Stage

- This first stage is a "Control Function" method: instead of instrumenting the endogenous regressors, we include additional regressors that capture the endogenous part of the error term.
- We are controlling for endogeneity by including $(\ell_{i,t-1}, k_{it}, i_{it})$ as "proxies" of ω_{it}.
- Key assumptions for the identification of α_L:

  (a) Invertibility of $f_K(\ell_{i,t-1}, k_{it}, \omega_{it}, r_t)$ *w.r.t* ω_{it}.

  (b) $r_{it} = r_t$, *i.e., no cross-sectional variability in unobservables, other than* ω_{it}, *affecting investment.*

  (c) Given $(\ell_{i,t-1}, k_{it}, i_{it}, r_t)$, *labor* ℓ_{it} *still has sample variability.*

Olley and Pakes First Stage

- **Example (with parametric linear investment func.):**

$$\boxed{\begin{array}{rl} (PF) & y_{it} = \alpha_L \ \ell_{it} + \alpha_K \ k_{it} + \omega_{it} + e_{it} \\ \\ (Inverse\ ID) & \omega_{it} = \gamma_1 \ i_{it} + \gamma_2 \ \ell_{i,t-1} + \gamma_3 \ k_{it} + \gamma_4 \ r_{it} \end{array}}$$

- Then,

$$y_{it} = \alpha_L \ \ell_{it} + (\alpha_K + \gamma_3) \ k_{it} + \gamma_1 \ i_{it} + \gamma_2 \ \ell_{i,t-1} + (\gamma_4 r_{it} + e_{it})$$

- Note that ℓ_{it} is correlated with r_{it}. Therefore, we need $r_{it} = r_t$ and include time dummies to control for r_t in order to have consistency of the OLS estimator in this regression.
- Note also that to identify ℓ_{it} with enough precision we need not high collinearity between this variable and $(k_{it}, i_{it}, \ell_{i,t-1})$.

Olley and Pakes Second Stage

- **Estimation of** α_K. It is based on the other two assumptions:

$$\begin{array}{ll} (OP-3) & \omega_{it} \text{ follows a first order Markov process.} \\ (OP-4) & k_{it} \text{ is decided at } t-1\text{: } k_{it} = (1-\delta) k_{i,t-1} + i_{i,t-1} \end{array}$$

- Since ω_{it} is first order Markov, we can write:

$$\omega_{it} = E[\omega_{it} \mid \omega_{i,t-1}] + \xi_{it} = h\left(\omega_{i,t-1}\right) + \xi_{it}$$

where ξ_{it} is an innovation which is mean independent of any information at $t-1$ or before. And $h(.)$ is some unknown function.

- ϕ_{it} is identified from 1st step; and $\phi_{it} = \alpha_K\ k_{it} + \omega_{it}$. Then,

$$\phi_{it} = \alpha_K\ k_{it} + h\left(\phi_{i,t-1} - \alpha_K\ k_{i,t-1}\right) + \xi_{it}$$

Olley and Pakes   Second Stage

- We estimate $h(.)$ and α_K by applying recursively the same type of semiparametric method as in the first stage of OP.

$$\phi_{it} = \alpha_K \ k_{it} + h\left(\phi_{i,t-1} - \alpha_K \ k_{i,t-1}\right) + \xi_{it}$$

- Suppose that we consider a quadratic function for $h(.)$: i.e., $h(\omega) = \pi_1\omega + \pi_2\omega^2$. Then:

$$\phi_{it} = \alpha_K \ k_{it} + \pi_1 \left(\phi_{i,t-1} - \alpha_K \ k_{i,t-1}\right) + \pi_2 \left(\phi_{i,t-1} - \alpha_K \ k_{i,t-1}\right)^2 + \xi_{it}$$

- It is clear that α_K, π_1 and π_2 are identified in this equation.

Olley and Pakes Second Stage

- Time-to build is a key assumption for the consistency of this method. If investment at period t is productive, then the equation becomes:

$$\phi_{it} = \alpha_K \ k_{i,t+1} + h\left(\phi_{i,t-1} - \alpha_K \ k_{it}\right) + \xi_{it}$$

- $k_{i,t+1}$ depends on investment at period t and therefore it is correlated with the innovation ξ_{it}.

OP: Empirical Application

- US Telecom. equipment industry: 1974-1987.
- Technological change and deregulation.
  - Elimination of barriers to entry;
  - Antitrust decisions against AT&T: The Consent Decree (implemented in 1984) $->$ divestiture of AT&T.
  - Substantial entry/exit of plants.
- Data: US Census of manufacturers.

OP: Empirical Application

TABLE VI

ALTERNATIVE ESTIMATES OF PRODUCTION FUNCTION PARAMETERS[a]
(STANDARD ERRORS IN PARENTHESES)

Sample:	Balanced Panel		Full Sample[c, d]						
								Nonparametric F_ω	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation Procedure	Total	Within	Total	Within	OLS	Only P	Only h	Series	Kernel
Labor	.851	.728	.693	.629	.628			.608	
	(.039)	(.049)	(.019)	(.026)	(.020)			(.027)	
Capital	.173	.067	.304	.150	.219	.355	.339	.342	.355
	(.034)	(.049)	(.018)	(.026)	(.018)	(.02)	(.03)	(.035)	(.058)
Age	.002	−.006	−.0046	−.008	−.001	−.003	.000	−.001	.010
	(.003)	(.016)	(.0026)	(.017)	(.002)	(.002)	(.004)	(.004)	(.013)
Time	.024	.042	.016	.026	.012	.034	.011	.044	.020
	(.006)	(.017)	(.004)	(.017)	(.004)	(.005)	(.01)	(.019)	(.046)
Investment	—	—	—	—	.13	—	—	—	—
					(.01)				
Other Variables	—	—	—	—	—	Powers of P	Powers of h	Full Polynomial in P and h	Kernel in P and h
# Obs.[b]	896	896	2592	2592	2592	1758	1758	1758	1758

OP: Empirical Application

- Going from OLS balanced panel to OLS full sample almost doubles α_K and reduces α_L by 20%. [Importance of endogenous exit].
- Controlling for simultaneity further increases α_K and reduces α_L.

OP: Empirical Application

TABLE XI

DECOMPOSITION OF PRODUCTIVITY[a]
(EQUATION (16))

Year	p_t	$\bar{p}_t$	$\Sigma_t \Delta s_{it} \Delta p_{it}$	$\rho(p_t, k_t)$
1974	1.00	0.90	0.01	−0.07
1975	0.72	0.66	0.06	−0.11
1976	0.77	0.69	0.07	−0.12
1977	0.75	0.72	0.03	−0.09
1978	0.92	0.80	0.12	−0.05
1979	0.95	0.84	0.12	−0.05
1980	1.12	0.84	0.28	−0.02
1981	1.11	0.76	0.35	0.02
1982	1.08	0.77	0.31	−0.01
1983	0.84	0.76	0.08	−0.07
1984	0.90	0.83	0.07	−0.09
1985	0.99	0.72	0.26	0.02
1986	0.92	0.72	0.20	0.03
1987	0.97	0.66	0.32	0.10

8. Control Function: Levinsohn-Petrin estimator

Levinshon & Petrin (2003)

- The main difference with OP method is that LP use the demand function for intermediate inputs instead of the investment equation to invert out unobserved productivity.

- Two main motivations:
  - Investment can be responsive to more persistent shocks in TFP; materials is responsive to every shock in TFP.
  - In some datasets **Zero Investment** accounts for a large fraction of the data. At $i_{it} = 0$ (corner solution / extensive margin) there is not invertibility between i_{it} and ω_{it}. Problems: loss of efficiency; missing estimates of TFP for many observations.

Levinshon & Petrin (2003)

- They consider a Cobb-Douglas production function in terms of labor, capital, and intermediate inputs (materials):

$$y_{it} = \alpha_L \ \ell_{it} + \alpha_K \ k_{it} + \alpha_M \ m_{it} + \omega_{it} + e_{it}$$

- Investment equation is replaced with demand for materials:

$$m_{it} = f_M \left(\ell_{i,t-1}, k_{it}, \omega_{it}, r_{it} \right)$$

- **Assumption LP-1**: $f_M \left(\ell_{i,t-1}, k_{it}, \omega_{it}, r_{it} \right)$ is invertible in ω_{it}.

- They maintain OP-2 [No other unobservables; $r_{it} = r_t$], OP-3 [Markov TFP], and OP-4 [Time-to-build].

Levinshon & Petrin: First Step

- Least squares estimation of parameter α_L and the nonparametric functions $\{\phi_t(.) : t = 1, 2, ..., T\}$ in regression equation:

$$y_{it} \quad = \quad \alpha_L \; \ell_{it} + \phi_t(\ell_{i,t-1}, k_{it}, m_{it}) + e_{it}$$

- $\phi_t(\ell_{i,t-1}, k_{it}, m_{it}) = \alpha_K \; k_{it} + \alpha_M m_{it} + f_M^{-1}(\ell_{i,t-1}, k_{it}, m_{it}, r_t)$ and f_M^{-1} is the inverse function of f_M with respect to ω_{it}.

Levinshon & Petrin: Second Step

- The second step is also similar to OP's second step but in the model with the intermediate input.
- ϕ_{it} is estimated in 1st step; and $\phi_{it} = \alpha_K \ k_{it} + \alpha_M \ m_{it} + \omega_{it}$. Then,

$$\phi_{it} \quad = \quad \alpha_K \ k_{it} + \alpha_M \ m_{it} + h\left(\phi_{i,t-1} - \alpha_K \ k_{i,t-1} - \alpha_M \ m_{i,t-1}\right) + \xi_{it}$$

- Important difference with OP: In this second step $E(m_{it} \ \xi_{it}) \neq 0$, i.e., materials m_{it} is endogenous.
- LP propose two approaches:
  - **"unrestricted method"**: instrument m_{it} with its lagged values [see GNR (2013) criticism];
  - **"restricted method"**: under static input, price-taking: $\alpha_M =$ Cost of materials/Revenue.

LP: Empirical application

- Plant-level data from 8 different Chilean manufacturing industries: 1979-1985 [Pinochet period].

LP: Empirical Application. Var input shares

TABLE 3

Average Nominal Revenue Shares (Percentage), 1979-85

Industry	Unskilled	Skilled	Materials	Fuels	Electricity
Metals	15.2	8.3	44.9	1.6	1.7
Textiles	13.8	6.0	48.2	1.0	1.6
Food Products	12.1	3.5	60.3	2.1	1.3
Beverages	11.3	6.8	45.6	1.8	1.5
Other Chemicals	18.9	10.1	37.8	1.7	0.7
Printing & Pub.	19.8	10.7	40.1	0.5	1.3
Wood Products	20.6	5.3	47.0	3.0	2.4
Apparel	14.0	4.9	52.4	0.9	0.3

LP: Empirical Application: Zeroes

TABLE 2

Percent of Usable Observations, 1979-85

Industry	Investment	Fuels	Materials	Electricity
Metals	44.8	63.1	99.9	96.5
Textiles	41.2	51.2	99.9	97.0
Food Products	42.7	78.0	99.8	88.3
Beverages	44.0	73.9	99.8	94.1
Other Chemicals	65.3	78.4	100	96.5
Printing & Pub.	39.0	46.4	99.9	96.8
Wood Products	35.9	59.3	99.7	93.8
Apparel	35.2	34.5	99.9	97.2

LP: Empirical Application: Zeroes

TABLE 4
Unrestricted and Restricted Parameter Estimates for 8 Industries
(Bootstrapped Standard Errors in Parentheses)

	Industry (ISIC Code)							
Input	311	381	321	331	352	322	342	313
Unskilled labor								
	0.138	0.164	0.138	0.206	0.137	0.163	0.192	0.087
	(0.010)	(0.032)	(0.027)	(0.035)	(0.039)	(0.044)	(0.048)	(0.082)
Skilled labor								
	0.053	0.185	0.139	0.136	0.254	0.125	0.161	0.164
	(0.008)	(0.017)	(0.030)	(0.032)	(0.036)	(0.038)	(0.036)	(0.087)
Materials								
	0.703	0.587	0.679	0.617	0.567	0.621	0.483	0.626
	(0.013)	(0.017)	(0.019)	(0.022)	(0.045)	(0.020)	(0.028)	(0.075)
Fuels								
	0.023	0.024	0.041	0.018	0.004	0.0162	0.053	0.087
	(0.004)	(0.008)	(0.012)	(0.018)	(0.020)	(0.016)	(0.014)	(0.027)
Capital								
unrestricted	0.13	0.09	0.08	0.18	0.17	0.10	0.21	0.08
	(0.032)	(0.027)	(0.054)	(0.029)	(0.034)	(0.024)	(0.042)	(0.050)
restricted	0.14	0.09	0.06	0.11	0.15	0.09	0.21	0.07
	(0.011)	(0.02)	(0.019)	(0.025)	(0.034)	(0.039)	(0.045)	(0.11)
Electricity								
unrestricted	0.038	0.020	0.017	0.032	0.017	0.022	0.020	0.012
	(0.021)	(0.010)	(0.024)	(0.028)	(0.032)	(0.014)	(0.024)	(0.022)
restricted	0.011	0.015	0.014	0.021	0.005	0.008	0.011	0.012
No. Obs.	6051	1394	1129	1032	758	674	507	465

8. Using First Order Conditions

First order conditions for flexible inputs

- Suppose that labor is a perfectly flexible input and the firm is a price-taker in output and labor markets. Then, F.O.C. imply:

$$P_{it} \ \frac{\partial Y_{it}}{\partial L_{it}} = W_{it}$$

- For the Cobb-Douglas PF, this condition becomes:

$$\alpha_L = \frac{W_{it} \ L_{it}}{P_{it} \ Y_{it}}$$

i.e., α_L is identified by the wage bill-to-revenue ratio.

- In fact, this condition rejects this simple version of the model. Substantial sample variation in $\frac{W_{it} \ L_{it}}{P_{it} \ Y_{it}}$. Either $\alpha_{L,it}$, or unobserved heterogeneity in cost of labor, or other assumptions do not hold.