Estimating Production Functions

Charlie Murry

Boston College

January 14, 2019

Motivation[1]

The production function is an object of primary economic interest.

- Characterize firm's cost function.
- What are returns to scale in an industry?
- How persistent is productivity?
- Do input elasticities change over time?
- What is the impact of new technology?
- What is impact of trade/macro policies on production?
- Is productivity correlated with exporting, importing, etc?

[1] I borrow from a variety of sources including Victor A's handbook chapter and Paul Grieco's and Peter Newberry's notes from Penn State and Paul Scott's notes from NYU.

More questions...

- What contributes to economic growth?
- Role of skill-biased tech change.
- Learning-by-doing?
- Do more productive firms choose higher quality or charge higher markups?

Fix Ideas

$$Y_{it} = L_{it}^{\beta_\ell} K_{it}^{\beta_k} U_{it} \tag{1}$$

- U – (total factor) productivity.
  - in contrast to labor productivity: Y/L
- L, K – input choices of firms.
- β – output elasticities.

Goal: Estimate β.

The catch: Choice of L a function of U.

Fix Ideas – Data

Typical dataset:

- Yearly (typically) data.
- Measure of output: quantity, revenues, value-added (R minus cost of materials)
- Inputs: laborers, capital stock, investment, R&D, materials, energy costs.
- USA: Compustat, US Census Long. Research Database.
- Europe: Panel data from central banks.
- Other: Chile, Columbia, France, Estonia, etc.

Framework: Production by Firms/plants

- Production is dynamic by its nature.

- Why? Capital (K) requires investment (I) and depreciates.

- Firms engage in other activities that are state variables.
  - R&D
  - Entry/exit
  - entry into new/export markets

- Unlike in demand estimation, it will be harder to ignore this for estimation because dynamics are at the core of even the baseline model.

- Models of industry dynamics rely on heterogeneous (and stochastic) productivity.
  - Jovanovic (1982), Hopanhyan (1992), Melitz (2003)

Econometric Simultaneity

Take logs of Eq 1:

$$y_{it} = \beta_\ell \ell_{it} + \beta_k k_{it} + \omega_{it} + \varepsilon_{it} \tag{2}$$

- ε is just random and not known at the time of choosing K and L.
- ω_{it} is productivity, known to the firm.
- ℓ_t or k_t could be chosen with knowledge of ω_t
- Typically we think that k is chosen at $t-1$ (as investment, i_{t-1})
- But more productive firms should choose more ℓ.

Classic solutions to the simultaneity problem

Input prices (or other variables) as instruments:

- Sometimes data not available.
- Almost never any variation in data across firms.
- Even if this exists, the variation is probably not exogenous.

Fixed Effects:

- Assumption: ω_{it} is really $\omega_i(+\varepsilon_{it})$
- Probably not suitable for long panels.
- Assumes no change in firm-specific technology – this rules out some interesting questions from the beginning.
- Differencing kills variation in k.

Classic solutions to the simultaneity problem

Dynamic Panel Methods

- See Arellano and Bond (1991).
- Let's put some structure on TFP.
- $\omega_t = \rho\omega_{t-1} + \xi_t$
- See problem set...

Econometric Selection

- Firms that exit have low productivity draws.
  - See Hopanhyan (1992) or Melitz (1993).
- Or, variable like “export status” will be endogenous, as more productive firms may be more willing to export/import.
- We will basically punt on this for now, but it is discussed at length in OP.

Biased Estimates

Simultaneity

- Upward bias in labor coefficient.
- More efficient firms choose more labor.

Selection

- Downward bias in capital coefficient.
- Firms with high capital have very low ω cutoffs for exit.
- exit distorts the joint dist. of k and ω.
- *Conditional on firm surviving*, there is negative correlation between k and ω, although in reality we expect low ω firms to be low k firms.

Roadmap of Talk

Olley and Pakes (1996)

Propose alternative solution to simultaneity (and selection) problem.

Idea

- Greater (optimal) investment this period implies the firm realized a higher ω_t.
- Use investment (i_t) to "proxy" for TFP (ω_t).
    - Remember: investment does not get used until next period!
- This only works under certain conditions – *which is the innovation in this paper.*

OP - Details[2]

OP1: Information

Firm info set ($\mathcal{I}_t$) includes only current and past ω. ε is exogenous: $E[\varepsilon_t \mid \mathcal{I}_t] = 0$.

OP2: First Order Markov

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

OP3: Timing

Firms accumulate capital according to $k_{it} = \kappa(k_{i,t-1}, i_{i,t-1})$. Labor is not dynamic.

OP4: Scalar Unobservable

Firms investment decisions: $i_{it} = f_t(k_{it}, \omega_{it})$.

OP5: Strict Monotonicity

$f_t(k_{it}, \omega_{it})$is strictly increasing in ω_{it}.

[2]The description here follows ACF.

Investment as a Proxy

OP2: First Order Markov
$p(\omega_{i,t+1} \mid I_{it}) = p(\omega_{i,t+1} \mid \omega_{it})$

OP5: Strict Monotonicity
$f_t(k_{it}, \omega_{it})$is strictly increasing in ω_{it}.

These together imply: firms with higher ω_{it} have higher expected marginal product of capital in the future and will invest more.

As ACF point out, proving this is tedious and may not extend to all models, but is general enough for relatively straightforward models with C-D or trans-log PFs.

- Involves analyzing the specific dynamic programming problem.

Econometric Procedure – First Stage

- Invert investment policy function:

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

  (remember we can do this because i is monotonic in ω!)

- Then plug in for ω in the linear (in logs) production function:

$$y_{it} = \beta_\ell \ell_{it} + \beta_k k_{it} + h(k_{it}, i_{it}) + \varepsilon_{it},$$

  or

$$y_{it} = \beta_\ell \ell_{it} + \Phi_t(k_{it}, i_{it}) + \varepsilon_{it}.$$

- (pause here and think about what is going on...)

Econometric Procedure – First Stage

- First stage regression:

$$y_{it} = \beta_\ell \ell_{it} + \Phi_t(k_{it}, i_{it}) + \varepsilon_{it},$$

  and recall

$$\Phi_t(k_{it}, i_{it}) = \beta_k k_{it} + h(k_{it}, i_{it}).$$

- Notice that k enters multiple places, and $h()$ is unknown, so the only primitive we can recover is β_ℓ.

- But we can also recover $\hat{\Phi}(k_{it}, i_{it})$ – treat this non-parametrically.

Econometric Procedure – Second Stage

- We have $\hat{\Phi}(k_{it}, i_{it})$ and $\hat{\beta}_\ell \ell_{it}$ in hand from the first stage.
- Decompose ω into the expected part (given the state $\mathcal{I}$) plus an innovation:

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

- Now plug this into the production function:

$$y_{it} = \alpha \ell_{it} + \beta k_{it} + g(\Phi_{t-1}(k_{i,t-1}, i_{i,t-1}) - \beta k_{i,t-1}) + \xi_{it} + \varepsilon_{it}$$

Econometric Procedure – Second Stage

- Use the following moment:

$$E[\xi_{it} + \varepsilon_{it} \mid \mathcal{I}_{it}] = 0$$

or

$$E[y_{it} - \alpha \ell_{it} - \beta k_{it} - g(\Phi_{t-1}(k_{i,t-1}, i_{i,t-1}) - \beta k_{i,t-1}) \mid \mathcal{I}_{it}] = 0.$$

Procedure[3]

1. plug-in first-stage estimates: $\hat{\alpha}$ and $\hat{\Phi}$ and pick parametric form for $g()$.
2. guess β_k.
3. Estimate $g()$:

$$\phi_{it} = \beta_k k_{it} + \underbrace{g(\phi_{i,t-1} - \beta_k k_{i,t-1})}_{E[\omega_t | \mathcal{I}_t]} + \xi_{it}$$

4. Back out $\hat{\xi}(\beta_k)$
5. Evaluate $E[\xi | \mathcal{I}]$
6. go back to 2. until $E[\xi | \mathcal{I}] = 0$.

[3] if g is linear then you could use OLS.

OP Results

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

Application

- How does productivity change of telecom equipment manufacturers after massive deregulation in telecom industry? [1977 and 1984]

- Productivity:

$$p_{it} = exp(y_{it} - \hat{\beta}_\ell \ell_{it} - \hat{\beta}_k k_{it})$$

TABLE IX

INDUSTRY PRODUCTIVITY GROWTH RATES[a]

Time Period	(1) Full Sample	(2) Balanced Panel
1974–1975	−.279	−.174
1975–1977	.020	−.015
1978–1980	.146	.102
1981–1983	−.087	−.038
1984–1987	.041	.069
1974–1987	.008	.020
1975–1987	.032	.036
1978–1987	.034	.047

[a]The numbers in Table IX are annual averages over the various subperiods.

Roadmap of Talk

Olley and Pakes

Levinsohn and Petrin

Ackerberg, Caves, and Frazer

Markups: De Loecker and Warzynski (2012)

Levinsohn and Petrin (2003)

Key Insight

- Today's choice of intermediate inputs is correlated with today's productivity.

Why is the important?

- Investment proxy not valid for plants with zero investment.
- Literature has shown investment is "lumpy".
  - May not immediate respond to productivity shocks.
  - Potential adjustment issues or measurement issues.

LP Estimation

- First Stage:

$$y_{it} = \beta_\ell \ell_{it} + \beta_k k_{it} + \beta_m m_{it}) + \varepsilon_{it}$$

$$y_{it} = \beta_\ell \ell_{it} + \Phi_t(k_{it}, m_{it}) + \varepsilon_{it}$$

- Estimate $\hat{\beta}_l$ and $\hat{\Phi}()$ treating Φ non-parametrically.

- Second Stage:

$$E[\xi_{it} + \varepsilon_{it} \mid \mathcal{I}_{it}] = 0$$

  or

$$E[y_{it} - \beta_\ell \ell_{it} - \beta_k k_{it} - \beta_m m_{it} - \\ g(\Phi_{t-1}(k_{i,t-1}, m_{i,t-1}) - \beta_k k_{i,t-1} - \beta_m m_{i,t-1}) \mid \mathcal{I}_{it}] = 0.$$

Discussion

1. We need a similar condition materials use is monotonically increasing in productivity, $m_t(\omega_t, k_t)$.
   - This relies on properties of the production function – in OP this relied on the Markov Perfect Eqm assumption.

   - Why? materials is a static input.

2. OP rules out firm specific unobservables affecting investment demand – firm-specific adjustment costs, prices, etc.
   - Why? The scalar unobservable assumption.

   - LP can allow for this because m is static, so $m(k_t, \omega_t)$ does not depend on investment.

3. In the second stage the need to identify two parameters so they use two moments:

$$E[\xi_t \begin{pmatrix} k_t \\ m_{t-1} \end{pmatrix}] = 0$$

Investment Zeros

TABLE 2

Per cent of non-zero observations

Industry (ISIC)	Investment	Fuels	Materials	Electricity
Food products (311)	42·7	78·0	99·8	88·3
Metals (381)	44·8	63·1	99·9	96·5
Textiles (321)	41·2	51·2	99·9	97·0
Wood products (331)	35·9	59·3	99·7	93·8

Results

	Industry (ISIC code)			
Input	311	381	321	331
Unskilled labour				
	0·139	0·172	0·130	0·193
	(0·010)	(0·033)	(0·024)	(0·034)
Skilled labour				
	0·051	0·188	0·155	0·133
	(0·009)	(0·025)	(0·026)	(0·030)
Electricity				
	0·085	0·081	0·005	0·047
	(0·007)	(0·015)	(0·019)	(0·021)
Fuels				
	0·023	0·020	0·038	0·021
	(0·004)	(0·011)	(0·010)	(0·014)
Materials				
	0·500	0·420	0·500	0·550
	(0·078)	(0·091)	(0·118)	(0·086)
Capital				
	0·240	0·290	0·180	0·190
	(0·053)	(0·094)	(0·095)	(0·090)
Returns to scale				
	1·037	1·172	1·007	1·133
	(0·059)	(0·075)	(0·113)	(0·157)

Roadmap of Talk

Olley and Pakes

Levinsohn and Petrin

Ackerberg, Caves, and Frazer

Markups: De Loecker and Warzynski (2012)

Ackerberg, Caves, and Frazer (2015)

Conceptual issue with OP/LP: functional dependence.

- Conditional on k, ω (and m), labor is completely determined...oops.

Example

- FOC for m is:

$$\beta_m K_{it}^{\beta_k} L_{it}^{\beta_\ell} M_{it}^{\beta_m - 1} e^{\omega_{it}} = \frac{p_m}{p_y}$$

- then plug this into the production function:

$$y_{it} = ln(\frac{1}{\beta_m}) + ln(\frac{p_m}{p_y}) + m_{it} + \varepsilon_{it}$$

- But there is no β_ℓ in this equation! So a moment condition based on ε will not have any co-variation in the data to inform about β_ℓ

Functional Dependence

The point is more general than the example above.

- In general, if ℓ is *only* a function of k, m, and t, then β_ℓ will not be identified (see Robinsin 1988 for a discussion of partially linear models).

What to do?

- Think about data generating process for ℓ that is realistic and implies ℓ is not completely pinned down by $\{k, m, t\}$.

- ACF consider timing assumptions on the shocks and the choice of ℓ.

- Solution? Assume "optimization error" in ℓ.
  - No. "Optimization error "in other vars (ie m) means we cannot invert $\omega(k, m)$.

  - In fact, measurement error in general causes problems in this literature.

Alternative DGP?

- Assume ℓ set at time t
- m set at time $t - b$ with $0 < b < 1$ – so like a half period.

$$y_{it} = \beta_{ell}\ell_{it} + \beta_m m_{it} + \beta_k k_{it} + \omega_{i,t-b} + \eta_{it}$$

  with

$$p(\omega_{t-b} \mid \mathcal{I}_{t-1}) = p(\omega_{t-b} \mid \omega_{t-1})$$

- there is an unobserved i.i.d. shock to ell after materials are chosen and before ℓ is chosen.
- ℓ now has its own shock and it is not a state variable.
- This works, but playing with timing like this gets a bit absurd.

ACF Procedure

- Consider a value added production function:

$$y_{it} = \beta_{ell}\ell_{it} + \beta_k k_{it} + \omega_{i,t-b} + \eta_{it}$$

  with

$$k_{it} = \kappa(k_{i,t-1}, i_{i,t-1})$$

- investment is chosen in $t-1$ and labor (ℓ) is potentially dynamic – it can be chosen t, $t-1$, $t-b$.

- Input demand $m_{it} = \tilde{f}_t(k_t, \ell_t, \omega_t)$ is monotonic in ω.

- demand for m is *conditional* on ℓ ! This is different (and more general).

ACF Procedure – First Stage

- In first stage, just recover expected output:

$$y_{it} = \Phi_t(m_t, \ell_t, k_t) + \varepsilon_{it}$$

  where

$$\Phi_t(m_t, \ell_t, k_t) = \beta_\ell \ell_{it} + \beta_k k_{it} + \tilde{f}_t(k_t, \ell_t, \omega_t)$$

- Now we can predict productivity:

$$\omega_{it} = \hat{\Phi}_t - \beta_\ell \ell_{it} - \beta_k k_{it}$$

ACF Procedure – Second Stage

- Non-parameterically regress $\hat{\omega}_{it}(\beta_\ell, \beta_k)$ on $\hat{\omega}_{it-1}(\beta_\ell, \beta_k)$ to get the "innovations" to productivity (just like LP).

$$\xi_{it} = \hat{\omega}_{it}(\beta_\ell, \beta_k) - E[\hat{\omega}_{it}(\beta_\ell, \beta_k) \mid \hat{\omega}_{it-1}(\beta_\ell, \beta_k)]$$

- Estimation using the following moments:

$$\frac{1}{T}\frac{1}{N}\sum_t \sum_i \hat{\xi} \begin{pmatrix} k_{it} \\ \ell_{i,t-1} \end{pmatrix}$$

- Intuition? If ℓ is chosen after $t-1$ then ℓ_t will be correlated with ξ_t, then:
  - ℓ_{t-1} is chosen without knowledge of ξ (the innovation),
  - Certainly k_t chosen without knowledge of ξ (i_t).

Discussion

Estimation Procedure

- The fact that this is a procedure for a "value added" production function is important.
- Timing assumption on ℓ but this does not conflict with m.
- Do not estimate β_ℓ in first stage, so functional dependence issue is gone.

Discussion

Measurement

- Ideal data is when we have *units* for each variable. (why?)
- Generally this is not true for output or inputs other than L.
  - Monetary units of inputs/output are not comparable.
  - If prices are observed (not common), then we can get back to units, and also use prices as instruments (if exogenous).
- Firms face downward sloping output demand curves (or up-sloping input supply)?
  - If every firm does not face the exact same demand, then investment functions will be different, OP/LP not valid.
  - Some recent advances along this point – Klette and Griliches (1996), De Loecker (2011).
- Basically, from the very start (data and measurement), there are a lot of assumptions baked into the standard (OP/LP) techniques.

MONTE CARLO RESULTS[a]

	ACF				LP			
	β_l		β_k		β_l		β_k	
Meas. Error	Coef.	Std. Dev.	Coef.	Std. Dev.	Coef.	Std. Dev.	Coef.	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 β_l and β_k are 0.6 and 0.4, respectively. Standard deviations reported are of parameter estimates across the 1000 replications.

Monte Carlo Description

- DGP1: Firms face different wages and choose wage at $t - 0.5$.
  - Favorable to ACF. Inconsistent with LP.
- DGP2: Optimization in ℓ. And data on m is planned materials.
  - Consistent with LP, although timing a bit ad hoc.
- DGP3: DGP1 + DGP2
  - Inconsistent with ACF and LP.

Roadmap of Talk

Olley and Pakes

Levinsohn and Petrin

Ackerberg, Caves, and Frazer

Markups: De Loecker and Warzynski (2012)

De Loecker and Warzynski (2012)

- Derive expression for markups ($\mu = P/MC$) from cost minimization of firm.
- This expression a function output elasticity. (ding ding!)
- We can use our new-found ability to estimate production functions to estimate markups.

Main Idea

Assumption: variable inputs are set each period to minimize costs.

- Cost minimization Lagrangian:

$$\mathcal{L}(X_{it}^{1}, X_{it}^{2}, K_{it}, \lambda_{it}) = \sum_{v=1}^{V} P_{it}^{v} X_{it}^{v} + r_{it} K_{it} + \lambda_{it}(Q_{it} - Q_{t}(\cdot_{it}))$$

- where X^{v} is a flexible input, P^{v} input prices.

Main Idea

Assumption: variable inputs are set each period to minimize costs.

- Cost minimization Lagrangian:

$$\mathcal{L}(X_{it}^1, X_{it}^2, K_{it}, \lambda_{it}) = \sum_{v=1}^{V} P_{it}^v X_{it}^v + r_{it} K_{it} + \lambda_{it}(Q_{it} - Q_t(\cdot_{it}))$$

- where X^v is a flexible input, P^v input prices.

- FOC:

$$P_{it}^v = \lambda_{it} \frac{\partial Q_t()}{\partial X_{it}^v}$$

- Notice that λ is the shadow cost of the Q constraint binding, i.e. the marginal cost of production.

Deriving Markup Expression

- Multiply by X^v/Q:

$$\frac{\partial Q_t()}{\partial X^v_{it}} \frac{X^v_{it}}{Q_{it}} = \frac{1}{\lambda} \frac{P^v_{it} X^v_{it}}{Q_{it}}$$

Deriving Markup Expression

- Multiply by X^v/Q:

$$\frac{\partial Q_t()}{\partial X_{it}^v}\frac{X_{it}^v}{Q_{it}} = \frac{1}{\lambda}\frac{P_{it}^v X_{it}^v}{Q_{it}}$$

- Multiply and divide by output price, P_{it}:

$$\underbrace{\frac{\partial Q_t()}{\partial X_{it}^v}\frac{X_{it}^v}{Q_{it}}}_{output\ elas.} = \mu_{it}\underbrace{\frac{P_{it}^v X_{it}^v}{P_{it}Q_{it}}}_{rev.\ share}$$

- or markup is a ratio of output elas. to revenue share,

$$\mu_{it} = \theta_{it}^v/\alpha_{it}^v$$

.

Markup Expression Discussion

$$\mu_{it} = \theta_{it}^{v} / \alpha_{it}^{v}$$

.

- This actually dates back to at Hall (1988) who computes industry markups.
- DWL compute plant level markups using careful techniques to estimate production functions.

Cobb-Douglas Example

- Output elasticity is a constant,

$$\theta_{it}^{L} = \frac{\partial Q_t()}{\partial L_{it}} \frac{X_{it}^{L}}{Q_{it}} = \beta_{\ell}$$

- Markup is then, $\mu_{it} = \frac{\beta_{\ell}}{\alpha_{it}^{L}}$.

Cobb-Douglas v Translog

- The C-D markup expression is not very flexible.

- Does not depend on levels of Q or L. Small firms have same markup as big firms for the same *input shares* and same *productivity*.

Translog Production Function

- Main specification of DLW,

$$y_{it} = \beta_k k_{it} + \beta_\ell \ell_{it} + \beta_{\ell\ell} \ell_{it}^2 + \beta_{ll} k_{it}^2 + \beta_{k\ell} k_{it} \ell_{it} + \omega_{it} + \varepsilon_{it}$$

DLW in practice

- Follow LP, but condition on extra covariates in the materials demand function

$$m_i t = m_t(k_{it}, \omega_{it}, \mathbf{z}_{it})$$

- Crucial assumption is that $m(\cdot)$ is invertible in ω.
- $\mathbf{z}$ must control for everything relevant – no leftover productivity that could move around markups.
- $\mathbf{z}$ includes lagged prices, lagged inputs, and **exporter status**.

DLW Findings

TABLE 2—ESTIMATED MARKUPS

Methodology	Markup
Hall[a]	1.03 (0.004)
Klette[a]	1.12 (0.020)
Specification	
I (Cobb-Douglas)	1.17
II (I w/ endog. productivity)	1.10
III (I w/ additional moments)	1.23
IV (Translog)	1.28
V (II w/ export input)	1.23
VI (Gross Output: labor)	1.26
VI (Gross Output: materials)	1.22
VII[a] (**I** w/ single markup)	1.16 (0.006)
VIII[a] (First difference)	1.11 (0.007)

[a]Markups are estimated jointly with the production function (as discussed in Section III), and we report the standard errors in parentheses. The standard deviation around the markups in specifications **I–VI** is about 0.5.

DLW Findings

TABLE 3—MARKUPS AND EXPORT STATUS I: CROSS-SECTION

Methodology	Export Premium
Hall	0.0155 (0.010)
Klette	0.0500 (0.090)
Specification	
I (Cobb-Douglas)	0.1633 (0.017)
II (I w/ endog. productivity)	0.1608 (0.017)
IV (Translog)	0.1304 (0.014)
V (II w/ export input)	0.1829 (0.017)
VIII (First difference)	0.1263 (0.013)

Notes: Estimates are obtained after running equation (21) where the different specifications refer to the different markup estimates, and we convert the percentage markup difference into levels as discussed above. The standard errors under specifications **I–V** are obtained from a nonlinear combination of the relevant parameter estimates. All regressions include labor, capital, and full year and industry dummies as controls. Standard errors are in parentheses.

Some Criticisms

That we won't have too much time for...and there are numerous

- C-D, trans-log, Hicks-neutral productivity
  - The "demand estimation" literature has spent a lot of effort making their methodology flexible.

- Relatedly, DLW sell their paper as estimating markups without "specifying how firms compete in the product market."
  - For different levels of concentrations should we expect different $m()$?
  - Concentration variable in $\mathbf{z}$?
  - More generally should markups affect input demand?
  - Have data only on revenues rears its ugly head too b/c firms are not price takers...

$$\frac{\partial R}{\partial X^v} = \frac{\partial Q}{\partial X^v}(P + \frac{\partial P}{\partial Q})$$