Estimating Production Functions with Expectations Data

Agnes Norris Keiller [1] Áureo de Paula [2]
John Van Reenen [1,3]

[1]Programme on Innovation and Diffusion, LSE
[2]UCL, IFS and CeMMAP
[3]MIT

DSE @ UCL 2025

Motivation

- Production functions integral to many strands of research

$$y_{it} = f(k_{it}, l_{it}; \theta) + u_{it}$$

- Extensive literature has presented various estimation methods
  - First order conditions: e.g., Solow (1957); Hall (1988)
  - Dynamic panel IV: e.g., Chamberlain (1982); Blundell and Bond (2000)
  - Proxy variables (control functions): e.g., Olley and Pakes (1996); Levinsohn and Petrin (2003); Ackerberg, Caves, and Frazer (2015)
  - Recent contributions: e.g., Gandhi, Navarro, and Rivers (2020); De Roux, Eslava, Franco, and Verhoogen (2021); Li (2021); Demirer (2022)
- Surveys increasingly elicit firms' expectations about future inputs and outputs (e.g., MES, MOPS)

- **Can we improve on existing production function estimators using data on firms' expectations?**

Preview

Can we improve on existing production function estimators using data on firms' expectations?

- Theory

  Expectations data allow one to relax assumptions of optimal firm choices required by proxy variable estimators

- MC simulations (time allowing!)

  Our proposed estimator is robust to optimisation error in inputs, while other methods are not

- ~~UK data over 2017-2020~~

Today

Methodology: Model, Identification and Estimation

Monte Carlo

Discussion

Today

Methodology: Model, Identification and Estimation

Monte Carlo

Discussion

The object of interest

- Consider a general production function of the following form

$$y_{it} = f(k_{it}, l_{it}; \theta) + \omega_{it} + e_{it}$$

  where $x \equiv \ln(X)$

- $\omega_{it} \equiv$ idiosyncratic productivity known by the firm when deciding period t input and investment
- $e_{it} = \epsilon_{it} + \upsilon_{it} \equiv$ unanticipated mean-zero disturbances

  $\epsilon_{it} \equiv$ productivity shocks unknown by the firm when making period t decisions
  $\upsilon_{it} \equiv$ measurement error

Dynamics

- ω_{it} follows a Markov process

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

- Capital evolves according to

$$K_{it} = (1 - \delta)K_{it-1} + I_{it-1}$$

where δ = the depreciation rate; I_{it-1} = investment

Proxy Methods

- OP\LP\ACF leverage firm decisions (input demand) and their monotonicity in ω (conditional on observables):
  - OP: firms' investment policy $\rightarrow \omega = \Phi^{OP}(i_{it}, k_{it})$
  - LP: firms' material input policy $\rightarrow \omega = \Phi^{LP}(m_{it}, k_{it})$
  - ACF: firms' material input policy $\rightarrow \omega = \Phi^{ACF}(m_{it}, l_{it}, k_{it})$
  - Typically justified by a model of optimal firm decisions

- We assume nothing about the invertibility of firms' decisions.

Expectations

- Firms form expectations about $t+1$ production and inputs at the end of t conditional on $\Omega_{it} \supset \{K_{it}, L_{it}, I_{it}, \omega_{it}, K_{it+1}\}$

- If firms' expectations align with the true production technology

$$\begin{aligned} \mathbb{E}_{it}[y_{it+1}|\Omega_{it}] &= \int f(k_{it+1}, l_{it+1}; \theta) dF_{it}(l_{it+1}) \\ &+ \mathbb{E}_{it}[\omega_{it+1}|\Omega_{it}] + \mathbb{E}_{it}[e_{it+1}|\Omega_{it}] \\ &= \int f(k_{it+1}, l_{it+1}; \theta) dF_{it}(l_{it+1}) + g(\omega_{it}) \end{aligned} \tag{1}$$

  - $F_{it}(l_{it+1}) \equiv$ firm i's subjective prob distribution over next-period labour input
  - $\mathbb{E}_{it}[\xi_{it+1}|\Omega_{it}] = \mathbb{E}_{it}[e_{it+1}|\Omega_{it}] = 0$

Recovering ω_{it}

- Rearranging equation (1) for $g(\omega_{it})$ obtains

$$g(\omega_{it}) = \mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \int f(k_{it+1}, l_{it+1}; \theta) dF_{it}(l_{it+1}) \tag{2}$$

- Assuming the RHS of equation (2) is strictly increasing in ω_{it}

$$\begin{aligned} \omega_{it} &= g^{-1}\left(\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \int f(k_{it+1}, l_{it+1}; \theta) dF_{it}(l_{it+1})\right) \\ &\equiv \Psi\left(\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \int f(k_{it+1}, l_{it+1}; \theta) dF_{it}(l_{it+1})\right) \end{aligned}$$

Identification

- Combining

$$\omega_{it} = \Psi\left(\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \int f(k_{it+1}, l_{it+1}; \theta) dF_{it}(l_{it+1})\right)$$

with

$$y_{it} = f(k_{it}, l_{it}; \theta) + \omega_{it} + e_{it}$$

$\Rightarrow$ a moment condition we can use to recover θ

$$\begin{aligned} 0 &= \mathbb{E}[e_{it}|\Omega_{it}] = \mathbb{E}[y_{it} - f(k_{it}, l_{it}; \theta) - \omega_{it}|\Omega_{it}] \\ &= \mathbb{E}\left[y_{it} - f(k_{it}, l_{it}; \theta) - \Psi\left(\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \int f(k_{it+1}, l_{it+1}; \theta) dF_{it}(l_{it+1})\right) |\Omega_{it}\right] \end{aligned}$$

Example

Cobb-Douglas production and $\omega_{it} = \rho\omega_{it-1} + \xi_{it}$ implies

$$\begin{aligned} y_{it} &= \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + e_{it} \\ &= \frac{\rho - 1}{\rho}\beta_0 + \beta_k k_{it} + \beta_l l_{it} + \frac{1}{\rho}\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \frac{\beta_k}{\rho} k_{it+1} - \frac{\beta_l}{\rho}\mathbb{E}_{it}[l_{it+1}|\Omega_{it}] + e_{it} \end{aligned}$$

- If measurements on the variables in red are available, we can identify and consistently estimate all the parameters!
- For more general laws of motion, one can use results in Robinson (1988) for identification.
- ...but how does one get measurements on $\mathbb{E}_{it}[y_{it+1}|\Omega_{it}]$ and $\mathbb{E}_{it}[l_{it+1}|\Omega_{it}]$?

Yes, Virginia, there is a... Management and Expectations Survey!

Looking ahead to the 2021 calendar year, what is the approximate turnover you would anticipate for this business in the following scenarios?

Lowest turnover

Report to the nearest £1,000. For example, £1,357,689 would be reported as £1,358,000

£ 2,800,000

Low turnover

£ 4,200,000

Medium turnover

£ 5,000,000

High turnover

£ 6,300,000

Highest turnover

£ 7,500,000

For the approximate turnover values you have just given for 2021, how likely do you think each scenario is to occur?

Your answers should add up to 100%

Likelihood of lowest turnover occurring

5 %

Likelihood of low turnover occurring

10 %

Likelihood of medium turnover occurring

60 %

Likelihood of high turnover occurring

20 %

Likelihood of highest turnover occurring

5 %

- MES 2017: turnover, employment, capital expenditure and expenditure on energy, goods and services
- MES 2020: turnover, employment

- To get what we need
  - Convert scenario responses into 5 points on a CDF
  - Estimate the parameters of a lognormal distribution to fit these points via minimisation

Lognormal Distribution Fit

Table: Avg Abs Devs Between Reported and Fitted Probabilities

	Mean	p25	p50	p75	N
	Restaurants				
Turnover	0.032	0.020	0.029	0.039	462
Employment	0.032	0.020	0.029	0.038	462
	Electronics				
Turnover	0.028	0.019	0.026	0.035	472
Employment	0.026	0.018	0.025	0.034	472

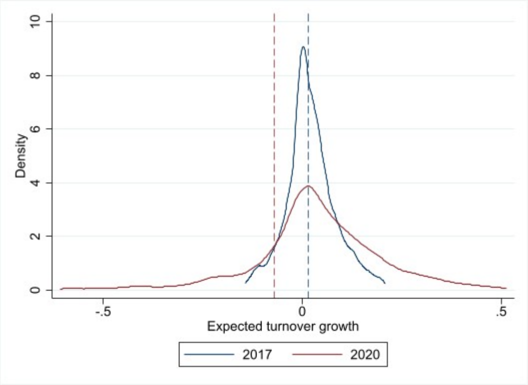
Lognormal Distribution Fit

Table: Median Fitted Subjective Distribution Characteristics

	μ	σ	Mean	S.D.	Median	IQR	N
	Restaurants						
Turnover	7.98	0.08	2848	1768	2811	321	462
Employment	4.27	0.09	67	41	67	8	462
	Electronics						
Turnover	8.73	0.07	6196	3750	6181	511	472
Employment	3.89	0.04	48	29	48	3	472

Turnover expectations far more pessimistic in 2020 (Covid effect?)

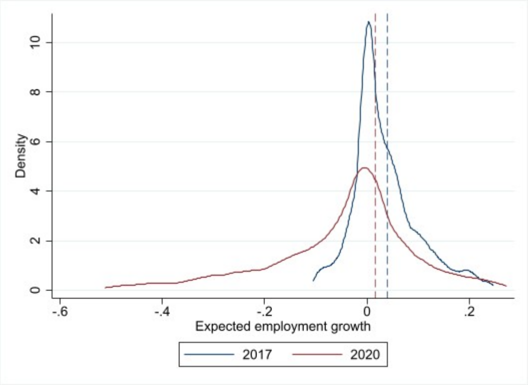
Figure: Expected turnover growth



Note: distribution is trimmed at the top and bottom 5%. Dashed lines denote means calculated across the entire distribution of 1.5% in 2017 and -7.1% in 2020. Sample sizes are 4388 in 2017 and 3127 in 2020.

Employment expectations slightly more pessimistic in 2020

Figure: Expected employment growth



Note: distribution is trimmed at the top and bottom 5%. Dashed lines denote means calculated across the entire distribution of 4.0% in 2017 and 1.7% in 2020. Sample sizes are 4388 in 2017 and 3127 in 2020.

Predictions and Outcomes

Table: Log Deviation Between Expected Levels and Outcomes

	Mean	S.D.	p50	N obs.	N firms
	Restaurants				
Turnover	-0.18	0.48	-0.13	170	150
Employment	-0.04	0.44	0.03	170	150
	Electronics				
Turnover	-0.01	0.29	-0.04	221	205
Employment	0.02	0.18	0.02	221	205

Identification

Under a more general law of motion, Cobb-Douglas production implies

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \Psi\left(\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \beta_k k_{it+1} - \beta_l \mathbb{E}_{it}[l_{it+1}|\Omega_{it}]\right) + e_{it}$$

Assuming Ψ is smooth, this is a *generalised additive model* (see, e.g., Hastie and Tibshirani (1986)) or *partially linear model* (see, e.g., Robinson (1988)):

$$y_{it} = \beta^\top x_{it} + h(z_{it}) + e_{it}$$

where $\beta \equiv [\beta_k \ \beta_l]^\top$, $x_{it} \equiv [k_{it} \ l_{it}]^\top$, $z_{it} \equiv [\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] \ k_{it+1} \ \mathbb{E}_{it}[l_{it+1}|\Omega_{it}]]^\top$ and

$$h(z_{it}) \equiv \Psi(\beta^\top z_{it})$$

Identification

Notice that:

$$\mathbb{E}[y_{it}|z_{it}] = \beta^\top \mathbb{E}[x_{it}|z_{it}] + h(z_{it}) \Rightarrow y_{it} - \mathbb{E}[y_{it}|z_{it}] = \beta^\top (x_{it} - \mathbb{E}[x_{it}|z_{it}]) + e_{it}$$

$\Rightarrow \beta$ is identified as long as $\mathbb{E}\left[(x_{it} - \mathbb{E}[x_{it}|z_{it}])(x_{it} - \mathbb{E}[x_{it}|z_{it}])^\top\right]$ is non-singular.

[This condition] prevents any element of X from being a.s. perfectly predictable by Z in the least squares sense. (...) Notice that (nonlinear) functional relations among X elements are not ruled out. Notice also that identification may be possible even if X uniquely defines Z, when the converse is not true. Robinson (ECTA, 1988, p.940)

Once we have β the nonparametric regression of $y_{it} - \beta_k k_{it} - \beta_l l_{it}$ on $\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \beta_k k_{it+1} - \beta_l \mathbb{E}_{it}[l_{it+1}|\Omega_{it}]$ identifies h (and, consequently, Ψ and g).

Identification

This holds when $f(\cdot)$ is linear in parameters (e.g., Cobb-Douglas, translog).

More generally,

Assume that $g(\cdot)$ is strictly monotonic and $\mathbb{E}[\epsilon_{it}|\Omega_{it}] = 0$. Let $x_{it} = (k_{it}, l_{it})$ and $z_{it} = (\mathbb{E}_{it}[y_{it+1}|\Omega_{it}], k_{it+1}, F_{it}(\cdot))$ and denote by $\theta_0 = (\beta_0, g_0)$ the data generating parameters. Then, if

$$f(x_{it};\beta_0) - \mathbb{E}\big[f(x_{it};\beta_0)|z_{it}\big] \neq f(x_{it};\beta) - \mathbb{E}\big[f(x_{it};\beta)|z_{it}\big]$$

with positive probability for any $\beta \neq \beta_0$, the parameter vector $\theta_0 = (\beta_0, g_0)$ is identified.

Estimation

Under a more general law of motion, Cobb-Douglas production implies

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \Psi\left(\mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \beta_k k_{it+1} - \beta_l \mathbb{E}_{it}[l_{it+1}|\Omega_{it}]\right) + e_{it}$$

Assuming Ψ is smooth, this is a *generalised additive model* (see, e.g., Hastie and Tibshirani (1986)) or *partially linear model* (see, e.g., Robinson (1988))

Problem 1: we require Ψ to be monotonic

- Impose constraints on the 1st derivatives of the smooth function that comprises Ψ (Pya and Wood (2015))

Problem 2: Ψ's argument is a function of the linear parameters

- Use an iterative 'backfitting' algorithm (Friedman and Stuetzle (1981), Ichimura and Todd (2007))

Estimation Protocol

Adapting the Friedman and Stuetzle (1981) algorithm to our setting:

1. Initialise the parameter vector at $\hat{\theta}_0 = (\hat{\beta}_{00}, \hat{\beta}_{k0}, \hat{\beta}_{l0})$
2. For iteration j, calculate $Z_{ij} = \mathbb{E}_{it}[y_{it+1}|\Omega_{it}] - \hat{\beta}_{0j-1} - \hat{\beta}_{kj-1} k_{it+1} - \hat{\beta}_{lj-1}\mathbb{E}_{it}[l_{it+1}|\Omega_{it}]$
3. Fit the model $y_{it} = \beta_0 - \beta_k k_{it} - \beta_l l_{it} + \Psi\left(Z_{ij}\right) + \epsilon_{it} + \upsilon_{it}$ using the shape constrained estimator of Pya and Wood (2015) to obtain $\hat{\theta}_j$
4. Calculate the Euclidean distance between $\hat{\theta}_j$ and $\hat{\theta}_{j-1}$. If the distance is below some tolerance level, stop and treat $\hat{\theta}_j$ as the model's parameter estimates. If not then set $j \leftarrow j+1$ and repeat from step 2

For the remainder of these slides, this algorithm is referred to as npr

Today

Methodology: Model, Identification and Estimation

Monte Carlo

Discussion

Monte Carlo Setup

- We follow ACF Monte Carlo setup: Details
  - y a Leontief composite of m and a 'value added' function of l and k.
  - ω follows an AR(1) process.
  - Investment subject to a firm-specific convex adjustment cost.
  - Firms face a common, time-invariant wage cost.
  - Optimisation error in l (as in ACF) but also in m and i.
- Firms' optimal decisions and expectations have an analytical solution.
- Simulate 1000 firms over 100 periods, use data from last 10.
- Compare OLS\LP\ACF\NPR across various DGPs.

Optimisation error in l

	$\beta_l = 0.6$				$\beta_k = 0.4$				
	Mean	**Median**	**S.D.**	**MSE**	**Mean**	**Median**	**S.D.**	**MSE**	**N runs**
NPR	0.600	0.600	0.003	0.000	0.400	0.400	0.005	0.000	500
OLS	0.919	0.919	0.002	0.102	0.098	0.098	0.004	0.091	500
OP	0.840	0.840	0.004	0.057	0.161	0.161	0.008	0.057	500
LP	0.600	0.600	0.003	0.000	0.400	0.400	0.014	0.000	500
ACF	0.918	0.600	4.716	22.293	0.021	0.399	4.950	24.597	500
ACF $\mid(\beta_l, \beta_k \in (0,1))$	0.600	0.600	0.009	0.000	0.399	0.399	0.016	0.000	489

Number of replications given in the 'N runs' column. ACF results from initialization at $\beta_{l0} = 0.5$ and $\beta_{k0} = 0.5$.

Optimisation error in (l, m, i)

	$\beta_l = 0.6$				$\beta_k = 0.4$				
	Mean	**Median**	**S.D.**	**MSE**	**Mean**	**Median**	**S.D.**	**MSE**	**N runs**
NPR	0.600	0.600	0.003	0.000	0.400	0.400	0.005	0.000	500
OLS	0.920	0.919	0.002	0.102	0.098	0.098	0.004	0.091	500
OP	0.916	0.916	0.002	0.100	0.143	0.151	0.030	0.067	500
LP	0.815	0.815	0.004	0.046	0.222	0.146	0.108	0.043	500
ACF	0.590	0.676	1.186	1.404	0.457	0.361	1.321	1.746	500
ACF $\mid(\beta_l, \beta_k \in (0, 1))$	0.676	0.677	0.019	0.006	0.362	0.361	0.023	0.002	497

Number of replications given in the 'N runs' column. ACF results from initialization at $\beta_{l0} = 0.5$ and $\beta_{k0} = 0.5$.

Monte Carlo summary

1. NPR robust to optimisation errors
   - (Outperforms OLS\LP\ACF as optimisation error $\sigma \uparrow$)
2. When optimisation error on labour only, NPR, LP and ACF do well – though ACF is at times numerically unstable, see footnote 16 in Ackerberg, Caves, and Frazer (2015) or Kim, Luo, and Su (2019).
3. When optimisation errors affect multiple inputs, NPR remains adequate but other estimators deteriorate as they rely on input choices to proxy for ω.

Today

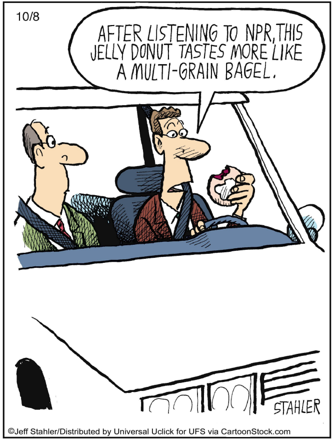
Methodology: Model, Identification and Estimation

Monte Carlo

Discussion

Discussion

- Generalisations (e.g., biased expectations, value added, imperfect knowledge of production function).
- Empirical application on UK data. Empirics
  - Different specifications (e.g., Cobb-Douglas, translog).
  - Analysis of implied productivity.
- Several other potential avenues to improve on this agenda though . . .
- Any suggestions welcome!


10/8
AFTER LISTENING TO NPR, THIS JELLY DONUT TASTES MORE LIKE A MULTI-GRAIN BAGEL.
STAHLER
©Jeff Stahler/Distributed by Universal Uclick for UFS via CartoonStock.com

The End!