Exploring ways to estimate endogenous productivity

Charlotte Guillard, Jordi Jaumandreu and
Jocelyn Olivari

CONCORDI Conference
Seville (Spain), 27-29 September 2018

Outline

- Motivation
- Description of the Spanish dataset
- Estimation results of the original DJ (2013) model **with** output prices
- Estimation results of the model **without** output prices

Motivation

- Productivity is **not** independent of the actions of firms (investments in R&D and innovation): assuming endogenous productivity gives more robust results

- Other paper have accounted for endogenous productivity in different contexts: Aw, Roberts & Xu (2011); De Loecker (2013); **Doraszelski & Jaumandreu (2013)**; Bøler, Moxnes & Ulltveit-Moe (2015); Jaumandreu & Yin (2016); Peters, Roberts, Vuong & Fryges (2016)

- The case of Latin America: Most countries of the region do not have data on output price.

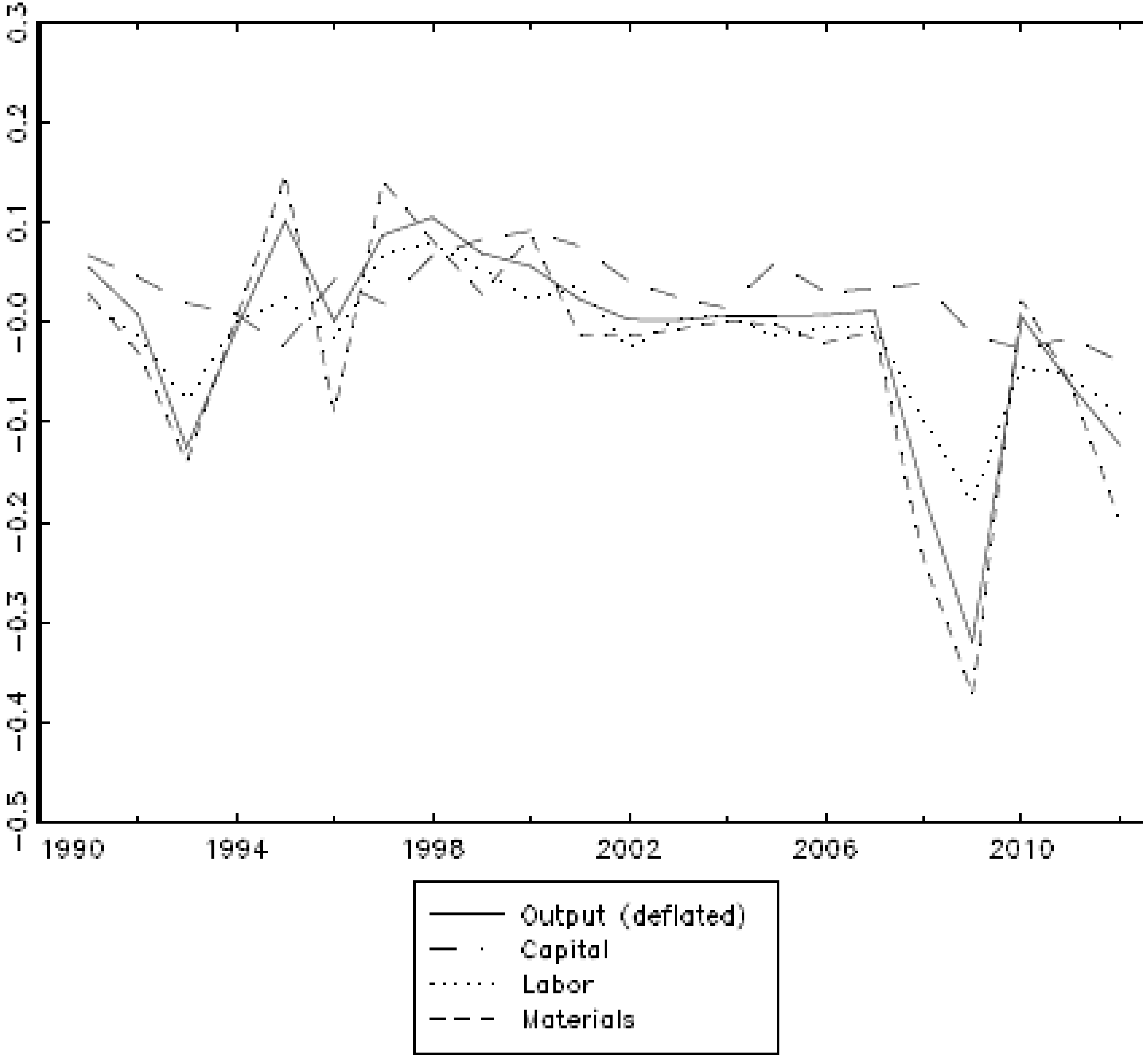
- Exploring ways to estimate productivity endogenously using ESEE Spanish dataset

Spanish dataset

- Panel dataset: Encuesta sobre Estrategias Empresariales (ESEE)
- Period: 1990-2012 (23 years)
- Representative sample of manufacturing firms: Firms with more than 200 workers are forcefully included. The rest is selected through stratified random sampling.

- 3.106 firms over the period
  - Only the firms with at least 3 consecutive years of observation are kept in the sample
- 20.799 observations over the period
  - 19737 (95%) observations correspond to firms with 200 employees and less
  - 9981 (48%) observations correspond to firms that realize R&D
  - 17292 (83%) observations correspond to exporting firms

9. Timber and furniture



Estimation model with output prices

- Depart from DJ (2013) with additional 13 years of data and slight modifications in the model

Cobb-Douglas production function (GMM)

$$y_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \boxed{\boldsymbol{\beta_t} + \boldsymbol{g}\left(\boldsymbol{\omega_{jt-1}}, \boldsymbol{a_{jt-1}}\right) + \boldsymbol{\xi_{jt}}} + e_{jt}$$

  - In-homogenous Markov process of order 1 to model endogenous productivity
  - Assume firms know their productivity at moment t and use this information to determine their demand for inputs to maximize their short run profits: inverse demand for input at t-1 $h_{jt-1}(.)$ to estimate ω_{jt-1}
  - η_{jt} is a function of output price and cycle (mdy)
  - To better capture capital and labor variation and improve coefficient estimation, there are two alternative ways:
    - Capacity utilization-adjusted capital: $k_{jt} = \tilde{k}_{jt} + cu_{jt}$
    - 2008 dummy interacted with capital and with labor in the production function: $\gamma_K d_{08} k_{jt} + \gamma_L d_{08} l_{jt}$

- Results
  - Reasonable, similar to DJ (2013)
  - Observe stochastic dominance in most industries

Estimation model without output prices

- In the absence of output price, we assume a demand for output:

$$Q_{jt} = \alpha_0 P_{jt}^{-\eta} e^{z_{jt}\alpha+\delta_{jt}}$$

Revenue

$$S_{jt} = \alpha_0 \boldsymbol{P}_{jt}^{-(\eta-1)} e^{z_{jt}\alpha+\delta_{jt}}$$

- Short-run profit maximization implies:

$$P_{jt} = \frac{\eta}{\eta-1} MC_{jt}$$

- Dual marginal cost:

$$MC_{jt} = \frac{VC_{jt}}{(\beta_L+\beta_M)F_{jt}} e^{-\omega_{jt}}\, e^{e_{jt}} = \overline{MC}_{jt} e^{-\omega_{jt}}\, e^{e_{jt}}$$

Replacing in the price by its optimal level in terms of marginal cost:

$$\boldsymbol{P}_{jt} = \frac{\eta}{\eta-1} \overline{MC}_{jt} e^{-\omega_{jt}}\, e^{e_{jt}}$$

Replacing in the revenue equation:

$$S_{jt} = \alpha_0 \left(\frac{\eta}{\eta-1}\overline{MC}_{jt}\right)^{-(\eta-1)} e^{z_{jt}\alpha+\boldsymbol{\delta_{jt}+(\eta-1)\omega_{jt}}-(\eta-1)e_{jt}}$$

Estimation model without output prices

Replacing in the revenue equation

$$S_{jt} = \alpha_0 \left(\frac{\eta}{\eta - 1} \overline{MC}_{jt} \right)^{-(\eta-1)} e^{z_{jt}\alpha + \boldsymbol{\delta_{jt}} + \boldsymbol{(\eta-1)\omega_{jt}} - (\eta-1)e_{jt}}$$

Assume composite unobservable $\tilde{\boldsymbol{\omega}}_{\boldsymbol{jt}}$ follows an in-homogenous Markov process:

$$\tilde{\boldsymbol{\omega}}_{\boldsymbol{jt}} = \beta_t + g(\tilde{\boldsymbol{\omega}}_{\boldsymbol{jt-1}}, r_{jt-1}) + \xi_{jt}$$

We use the inverse demand for labor at time t-1 $\boldsymbol{h_{jt-1}}(.)$ to estimate the unobservable composite term at time t-1:

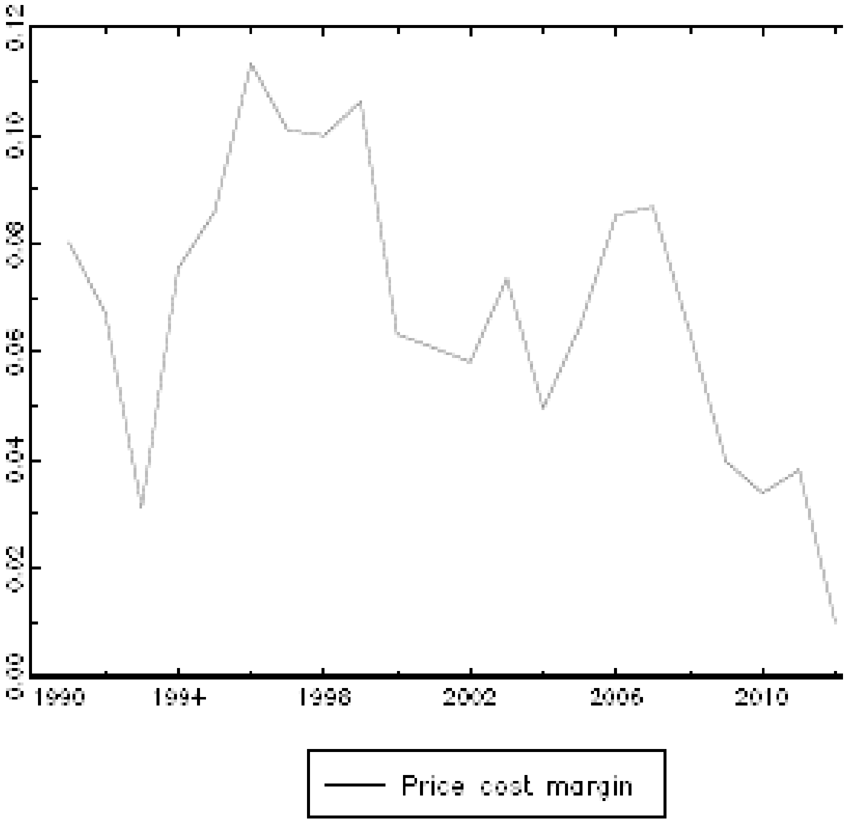
$$\tilde{\boldsymbol{\omega}}_{\boldsymbol{jt}} = \beta_t + g(\boldsymbol{h_{jt-1}}(.), r_{jt-1}) + \xi_{jt}$$

Estimation model without output prices

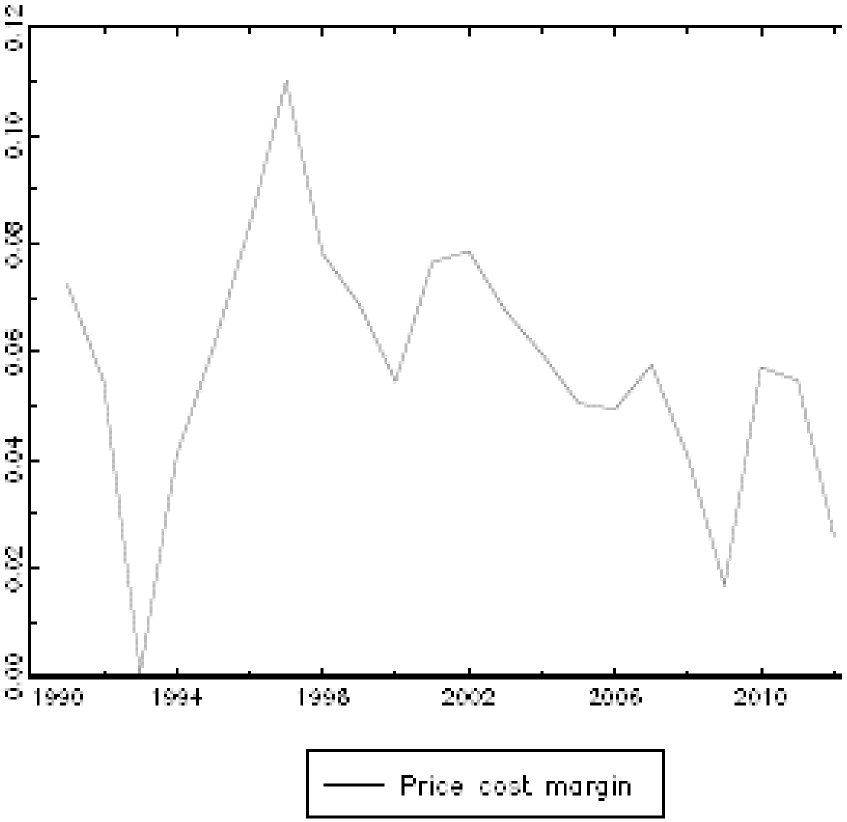
- (1) Revenue equation – in log (GMM)

$$s_{jt} = \ln \alpha_0 - (\eta_{jt} - 1)\ln\left(\frac{\eta_{jt}}{\eta_{jt} - 1}\right) - (\eta_{jt} - 1)\overline{mc}_{jt} + z_{jt}\alpha + \boldsymbol{\beta_t} + \boldsymbol{g}\left(\boldsymbol{h_{jt-1}}(.), \boldsymbol{r_{jt-1}}\right) + \boldsymbol{\xi_{jt}} + u_{jt}$$

5. Electrical goods



6. Transport equipement



$$\frac{P_{jt} - MC_{jt}}{P_{jt}} = -\frac{1}{\eta_{jt}}$$

Estimation model without output prices

- (1) Revenue equation – in log (GMM)

$$s_{jt} = \ln\alpha_0 - (\eta_{jt} - 1)\ln\left(\frac{\eta_{jt}}{\eta_{jt} - 1}\right) - (\eta_{jt} - 1)\overline{mc}_{jt} + z_{jt}\alpha + \boldsymbol{\beta_t} + \boldsymbol{g}\left(\boldsymbol{h_{jt-1}}(.), \boldsymbol{r_{jt-1}}\right) + \boldsymbol{\xi_{jt}} + u_{jt}$$

- (2) Gross margin equation to estimate υ and η_{jt} simultaneously (NLS)

$$gm_{jt} = s_{jt} - vc_{jt} = -ln(\upsilon) + ln\left(\frac{\eta_{jt}}{\eta_{jt-1}}\right) + e_{jt}$$

$$gm_{jt} = m_0 + m_1.mdy_{jt} + e_{jt} = a_{jt} + e_{jt} \quad \text{so that} \quad \eta_{jt} = \frac{\upsilon e^{a_{jt}}}{\upsilon e^{a_{jt}} - 1}$$

- Estimation of the system of equations in 2 steps

- Results
  - β_k around 0.1-0.2, $\upsilon \approx 1$, η between 6.5 and 24.5

Mean difference in productivity (R&D vs No-R&D)

MODEL WITHOUT PRICE (columns) / **MODEL WITH PRICE** (rows)

	Positive (4/20)	Negative (16/20)
Positive (11/20)	**5** (2/20) • Transport (big) • Timber & furniture (small)	**6** (9/20) • Chemicals (small & big) • Machinery (small & big) • Electrical (big) • Transport (small) • Food & tobacco (small & big) • Textile (big)
Negative (9/20)	**4** (2/20) • Electrical (small) • Paper (small)	**5** (7/20) • Metal (small & big) • NM Minerals (small & big) • Textile (small) • Timber & furniture (big) • Paper (big)

Stochastic dominance

MODEL WITHOUT PRICE

MODEL WITH PRICE

	Stochastic dominance (0/10)	Partial stochastic dominance (9/10)	No Stochastic dominance (1/10)
Stochastic dominance (5/10)		• Chemicals • Machinery • Transport • Timber & furniture	• Food & tobacco
Partial stochastic dominance (3/10)		• Non-metallic minerals • Textile • Paper	
No stochastic dominance (2/10)		• Metals • Electrical	

Estimation model without output prices

- Model is based on a composite that DOES NOT measure only productivity: it contains unobservable productivity, unobservable demand heterogeneity and demand elasticity (scaling factor).
- Important differences in the results of mean differences in productivity and stochastic dominance → Importance of unobserved demand heterogeneity (Foster, Haltiwanger & Syverson, 2008; Jaumandreu & Yin, 2016)
- Possible explanations?
  - R&D might not have the same effect on productivity in all sectors
  - If product differentiation is big in a sector, it has a demand advantage that can be translated into a cost disadvantage ultimately lowering productivity. → This seems to be the case of some firms in China (see Jaumandreu & Yin, 2016).

Final remarks and next steps

- Exploring endogenous productivity models with and without prices: composite productivity does not behave as productivity: affects the measure of returns to innovation efforts

- How to deal with this and estimate endogenous productivity in the absence of prices?
  - Estimation with groups of more homogeneous firms to limit variability
  - Separate unobserved productivity from unobserved demand heterogeneity by estimating using two markets (regional markets) as in Jaumandreu & Yin (2016)

- Next steps
  - Improving the estimations with better instruments and specifications (especially the margin equation)
  - Correcting for standard errors