Empirical estimation of non linear input-output models: an Entropy Econometrics approach

Esteban Fernández-Vázquez

University of Oviedo (Spain)
-Literature on non linear input output modeling is mainly theoretical (Lahiri, 1973; Lahiri and Pyatt, 1980; Chander, 1983; Fujimoto, 1986, or Dietzenbacher, 1994)

-Empirical estimation prevented by data availability: number of parameters to estimate higher than available data points

-CGE models: calibration techniques
- Proposal here: to estimate the parameters of non linear IO models

- Exploit new IO databases

- Econometric estimation based on Entropy Econometrics (EE), econometric technique suitable for ill-conditioned datasets
Outline of the presentation

1. A simple non linear IO model
2. Overview of entropy econometrics
3. Estimating non linear IO models with EE
4. An illustration: a non linear IO model for Spain
5. Final remarks
1. A simple non linear IO model

- The traditional linear IO model (1):

\[ z_{ij} = (a_{ij})x_j \]
\[ 0 \leq a_{ij} \leq 1 \]

\[ x = (I - A)^{-1}f = Lf \]

- A non-linear (scale-dependent) IO model (2):

\[ z_{ij}(x_j) = \alpha_{ij}x_j^{\beta_{ij}} \]
\[ \alpha_{ij}, \beta_{ij} \geq 0 \]

\[ x = A^*(x)x + f = L^*(x)f \]

- (2) is equivalent to (1) when \( \beta_{ij} = 1 \)
Multipliers with IO models

- A linear model:

\[ r = \hat{c}x = \hat{c}(I - A)^{-1}f = Mf \]

\[ c_j = r_j / x_j \]

Scale-independent coefficients

- A model with scale-dependent multipliers:

\[ r = \hat{c}^*(x)x = \hat{c}^*(x)(I - A^*(x))^{-1}f = M^*(x)f \]

\[ c_j^*(x_j) = r_j / x_j = \alpha^r_j x_j^{\beta^r_j} / x_j = \alpha^r_j x_j^{(\beta^r_j - 1)} \]

Use of variable \( r \) depends on the output

Parameters (\( \alpha, \beta \) and \( \alpha^r, \beta^r \)) in the equations condition the effect of final demand shocks
2. Overview of entropy econometrics

- a discrete random variable that can take $M \geq 2$ values $X = \{x_1, \ldots, x_M\}$ with probabilities $p = \{p_1, \ldots, p_M\}$. Shannon’s entropy function measures the uncertainty in $X$:

$$\max_p H(p) = -\sum_{m=1}^{M} p_m \ln(p_m)$$

- $H(p)$ achieves its unconstrained maximum for the uniform distribution. $H(p) = 0$ means no uncertainty

- if $p$ was unknown what would be our best guess? The distribution that maximizes $H(p)$. If we have some information about $X$, we now maximize $H(p)$ subject to (but only to) the information we have
- The underlying ideas of the ME methodology can be applied for estimating the parameters of general linear models (GME)

- The general form (T observations, H parameters) is:

\[ y = X\beta + \epsilon \]

- The beta parameters (and the error) do not behave as probabilities
-) each parameter $\beta_h$ is modeled as a discrete random variable that can take $M \geq 2$ values

-) define a support vector $b$ that contains the $M \geq 2$ possible outcomes of the random variable:

$$b' = (b_1, ..., b_M)$$

-) each element of $b'$ has an unknown probability:

$$p'_h = (p_{h1}, ..., p_{hM})$$

-) each parameter $\beta_h$ is given by the following expression (same procedure for the errors):

$$\beta_h = b' p_h = \sum_{m=1}^{M} b_m p_{hm}; \forall h = 1, ..., H$$
The GME program

\[
\text{Max}_{\mathbf{P}, \mathbf{U}} H(\mathbf{P}, \mathbf{U}) = - \sum_{h=1}^{H} \sum_{m=1}^{M} p_{hm} \ln(p_{hm}) - \sum_{n=1}^{N} \sum_{j=1}^{J} u_{tj} \ln(u_{tj})
\]

subject to:

\[
y_t = \sum_{k=1}^{K} \sum_{m=1}^{M} b_m p_{hm} x_{kt} + \sum_{j=1}^{J} v_j u_{tj}; \quad \forall t = 1, \ldots, T
\]

\[
\sum_{m=1}^{M} p_{hm} = 1; \quad \forall h = 1, \ldots, H
\]

\[
\sum_{j=1}^{J} u_{tj} = 1; \quad \forall t = 1, \ldots, T
\]

The probabilities are estimated by maximizing entropy \textbf{conditional on the observations}.
- once the $p$ probabilities are recovered, we have point-estimates of the parameters:

$$\hat{\beta}_h = \sum_{m=1}^{M} b_m \hat{p}_{hm}; \forall h = 1, ..., H$$

- under some mild assumptions, GME estimators are asymptotically consistent and normally distributed:

$$\hat{\beta} \rightarrow N[\beta, \hat{\sigma}^2(X'X)^{-1}]$$

t-ratio statistics can be calculated

$$\hat{\sigma}^2_h = \sigma^2_e \left( \frac{\sigma^2_b}{\sigma^2_b + \sigma^2_v} \right), \forall h = 1, ..., H;$$
3. Estimating non linear IO models with EE

- EE can be applied to recover the parameters of nonlinear IO equations. We need:

  - a dataset (a time series or a cross-section) of IO tables
  - to assume that the production technology is constant along all the observations

- \( T \) observations of a \((n \times n)\) IO matrix, \( n^2 \) equations like:

\[
\Delta \ln(z_{ijt}) = \Delta \beta_{ij} \ln(x_{jt}) + \varepsilon_{ijt}
\]

- Support vectors:

  - “natural” centers and bounds
  - 3-sigma rule for the errors, with \( J=3, \ v=(-3s,0,3s) \)
  - with \( M=3, \ b=(b_1,b_2,b_3)=(b_2-d,b_2,b_2+d)=(0,1,2) \)
A global GME program for \((n \times n)\) equations

\[ \text{Max } H(P_b, W) \]

subject to:

\[ \Delta \ln (Z) = BP \Delta \ln (X) + VW \]

\[ e' p_{ij} = 1; \forall i, j \]

Without any information the GME solution is the linear model (all \(\beta_{ij}=1\))

\[ e' w_{ijt} = 1; \forall i, j, t \]

\[ e' \hat{A}^*(x_t) < 1 \]

\[ [I - \hat{A}^*(x_t)]^{-1} = \left[ I + \hat{A}^*(x_t) + [\hat{A}^*(x_t)]^2 + [\hat{A}^*(x_t)]^3 + \cdots \right] \]
4. An illustration: a non linear IO model for Spain

-A non-linear IO model will be estimated for Spain in order to estimate output and labor multipliers

- 1995-2009 series of annual industry-by-industry IO tables (from WIOD)

- 1995-2009 series of annual labor figures by industry (from WIOD)

- aggregation into 16 sectors:

\[ z_{ijt}(x_{jt}) = \alpha_{ij} x_{it}^{\beta_{ij}} \quad 256 \text{ equations} \]

\[ r_{jt}(x_{jt}) = \alpha_j^r x_{jt}^{\beta_j^r} \quad 16 \text{ equations} \]
## Appendix: Industry classification

<table>
<thead>
<tr>
<th>Industry number</th>
<th>Industry description</th>
</tr>
</thead>
<tbody>
<tr>
<td>i1</td>
<td>Agriculture, Hunting, Forestry and Fishing</td>
</tr>
<tr>
<td>i2</td>
<td>Mining and Quarrying</td>
</tr>
<tr>
<td>i3</td>
<td>Food, Beverages and Tobacco</td>
</tr>
<tr>
<td>i4</td>
<td>Textiles, Leather and Footwear products</td>
</tr>
<tr>
<td>i5</td>
<td>Wood, pulp and paper</td>
</tr>
<tr>
<td>i6</td>
<td>Coke, Refined Petroleum and Nuclear Fuel</td>
</tr>
<tr>
<td>i7</td>
<td>Chemicals, rubber, plastics and non-metallic mineral</td>
</tr>
<tr>
<td>i8</td>
<td>Basic Metals and Fabricated Metal</td>
</tr>
<tr>
<td>i9</td>
<td>Machinery, equipment and n.e.c. manufacturing</td>
</tr>
<tr>
<td>i10</td>
<td>Electricity, Gas and Water Supply</td>
</tr>
<tr>
<td>i11</td>
<td>Construction</td>
</tr>
<tr>
<td>i12</td>
<td>Sale, maintenance and trade</td>
</tr>
<tr>
<td>i13</td>
<td>Hotels and Restaurants</td>
</tr>
<tr>
<td>i14</td>
<td>Transport, post and telecommunications</td>
</tr>
<tr>
<td>i15</td>
<td>FIRE services and other business activities</td>
</tr>
<tr>
<td>i16</td>
<td>Other services</td>
</tr>
</tbody>
</table>
- Equations are estimated in FD
- Point estimates and t-ratio statistics for the \((\alpha, \beta)\) parameters are computed
- t-ratios are used to test for linearity
  - \(H_0: \beta_{ij}=1\)
  - \(H_1: \beta_{ij} \neq 1\)
Table 1: GME estimates of the $\beta_{ij}$ parameters

<table>
<thead>
<tr>
<th>$\beta_{ij}$</th>
<th>i1</th>
<th>i2</th>
<th>i3</th>
<th>i4</th>
<th>i5</th>
<th>i6</th>
<th>i7</th>
<th>i8</th>
<th>i9</th>
<th>i10</th>
<th>i11</th>
<th>i12</th>
<th>i13</th>
<th>i14</th>
<th>i15</th>
<th>i16</th>
</tr>
</thead>
<tbody>
<tr>
<td>i1</td>
<td>1.067</td>
<td>0.914</td>
<td>0.980</td>
<td>0.916</td>
<td>0.990</td>
<td><strong>0.470</strong></td>
<td>1.089</td>
<td><strong>0.838</strong></td>
<td>0.949</td>
<td><strong>0.706</strong></td>
<td><strong>0.590</strong></td>
<td>0.874</td>
<td>0.865</td>
<td>0.943</td>
<td>0.889</td>
<td>0.517</td>
</tr>
<tr>
<td>i2</td>
<td>1.038</td>
<td><strong>1.117</strong></td>
<td>1.076</td>
<td>1.120</td>
<td>1.082</td>
<td>1.165</td>
<td>1.139</td>
<td>1.039</td>
<td>1.086</td>
<td><strong>0.641</strong></td>
<td>0.932</td>
<td>0.972</td>
<td>1.004</td>
<td>1.001</td>
<td>0.983</td>
<td>0.998</td>
</tr>
<tr>
<td>i3</td>
<td>0.944</td>
<td>0.894</td>
<td>0.957</td>
<td><strong>0.803</strong></td>
<td>*0.914</td>
<td><strong>0.561</strong></td>
<td><strong>0.911</strong></td>
<td>0.838</td>
<td><strong>0.704</strong></td>
<td><strong>0.911</strong></td>
<td><strong>0.780</strong></td>
<td>0.817</td>
<td>1.012</td>
<td>0.844</td>
<td>0.862</td>
<td>0.888</td>
</tr>
<tr>
<td>i4</td>
<td>1.006</td>
<td>0.987</td>
<td>1.026</td>
<td>1.057</td>
<td>1.060</td>
<td><strong>0.454</strong></td>
<td>0.990</td>
<td>1.026</td>
<td>0.960</td>
<td>0.919</td>
<td>1.022</td>
<td>0.975</td>
<td>0.953</td>
<td>1.015</td>
<td>0.981</td>
<td>0.989</td>
</tr>
<tr>
<td>i5</td>
<td>1.059</td>
<td>0.892</td>
<td><strong>0.897</strong></td>
<td><strong>0.971</strong></td>
<td>0.971</td>
<td>0.295</td>
<td>0.981</td>
<td>0.959</td>
<td>0.952</td>
<td>0.948</td>
<td>0.874</td>
<td>0.977</td>
<td>0.815</td>
<td><strong>0.797</strong></td>
<td>0.827</td>
<td>0.908</td>
</tr>
<tr>
<td>i6</td>
<td>0.771</td>
<td>0.922</td>
<td>1.048</td>
<td>1.098</td>
<td>1.023</td>
<td><strong>1.850</strong></td>
<td>0.910</td>
<td>0.976</td>
<td>0.944</td>
<td>0.942</td>
<td><strong>0.747</strong></td>
<td>0.949</td>
<td>0.865</td>
<td>0.904</td>
<td>0.872</td>
<td>0.881</td>
</tr>
<tr>
<td>i7</td>
<td>0.906</td>
<td><strong>0.829</strong></td>
<td><strong>0.927</strong></td>
<td>0.920</td>
<td>0.938</td>
<td><strong>0.513</strong></td>
<td>0.974</td>
<td><strong>0.886</strong></td>
<td>0.918</td>
<td>0.855</td>
<td><strong>0.891</strong></td>
<td>0.993</td>
<td><strong>0.510</strong></td>
<td>0.900</td>
<td>0.889</td>
<td>0.943</td>
</tr>
<tr>
<td>i8</td>
<td>1.027</td>
<td>1.003</td>
<td>0.967</td>
<td>1.093</td>
<td>1.033</td>
<td><strong>0.381</strong></td>
<td>1.088</td>
<td><strong>1.075</strong></td>
<td><strong>1.001</strong></td>
<td>0.971</td>
<td>0.951</td>
<td>1.073</td>
<td>0.850</td>
<td>0.978</td>
<td>0.974</td>
<td>0.945</td>
</tr>
<tr>
<td>i9</td>
<td>1.004</td>
<td>1.072</td>
<td>1.103</td>
<td>1.102</td>
<td>1.011</td>
<td><strong>0.556</strong></td>
<td><strong>1.059</strong></td>
<td>1.055</td>
<td>1.079</td>
<td>0.949</td>
<td>0.979</td>
<td>1.075</td>
<td>0.980</td>
<td>1.052</td>
<td>0.988</td>
<td>0.965</td>
</tr>
<tr>
<td>i10</td>
<td>0.989</td>
<td>0.991</td>
<td>1.103</td>
<td>1.041</td>
<td>1.032</td>
<td>0.874</td>
<td>1.073</td>
<td>1.026</td>
<td>0.868</td>
<td><strong>1.313</strong></td>
<td><strong>0.908</strong></td>
<td>1.058</td>
<td>1.071</td>
<td>0.963</td>
<td>1.023</td>
<td>1.061</td>
</tr>
<tr>
<td>i11</td>
<td>1.005</td>
<td>0.995</td>
<td>0.969</td>
<td>1.022</td>
<td>1.023</td>
<td><strong>0.489</strong></td>
<td>1.055</td>
<td>0.941</td>
<td>0.912</td>
<td>1.046</td>
<td>1.103</td>
<td>1.054</td>
<td>0.959</td>
<td>1.115</td>
<td>1.021</td>
<td>1.041</td>
</tr>
<tr>
<td>i12</td>
<td>0.999</td>
<td>0.914</td>
<td>1.013</td>
<td><strong>0.907</strong></td>
<td>1.050</td>
<td><strong>1.319</strong></td>
<td>0.994</td>
<td>0.943</td>
<td>0.890</td>
<td>0.910</td>
<td><strong>0.889</strong></td>
<td>0.982</td>
<td>1.044</td>
<td>1.051</td>
<td>0.951</td>
<td>1.031</td>
</tr>
<tr>
<td>i13</td>
<td>1.068</td>
<td>0.898</td>
<td>1.069</td>
<td>1.041</td>
<td>0.983</td>
<td><strong>0.338</strong></td>
<td>0.985</td>
<td>0.845</td>
<td>0.507</td>
<td>0.978</td>
<td>0.974</td>
<td>0.943</td>
<td>1.005</td>
<td>0.983</td>
<td>0.927</td>
<td>1.003</td>
</tr>
<tr>
<td>i14</td>
<td>1.086</td>
<td>1.005</td>
<td>1.071</td>
<td>0.824</td>
<td>0.974</td>
<td>1.254</td>
<td>0.977</td>
<td>0.932</td>
<td><strong>0.803</strong></td>
<td>1.115</td>
<td>0.935</td>
<td>1.045</td>
<td>1.100</td>
<td>1.055</td>
<td>0.933</td>
<td>1.032</td>
</tr>
<tr>
<td>i15</td>
<td>1.027</td>
<td>1.028</td>
<td>0.941</td>
<td>0.897</td>
<td>0.957</td>
<td>0.784</td>
<td>1.001</td>
<td>0.896</td>
<td>0.859</td>
<td>1.150</td>
<td>0.974</td>
<td><strong>1.073</strong></td>
<td>1.084</td>
<td>1.053</td>
<td>1.024</td>
<td>1.069</td>
</tr>
<tr>
<td>i16</td>
<td>0.990</td>
<td><strong>0.894</strong></td>
<td>1.066</td>
<td>0.936</td>
<td>1.051</td>
<td><strong>0.519</strong></td>
<td>1.000</td>
<td>0.888</td>
<td><strong>0.800</strong></td>
<td>1.052</td>
<td>0.987</td>
<td>0.988</td>
<td>1.082</td>
<td>1.098</td>
<td>1.008</td>
<td><strong>1.106</strong></td>
</tr>
</tbody>
</table>

** and * stands for estimates significantly different from 1 at 5% and 10% respectively
- Confidence intervals for multipliers can be computed to complement our point estimates.

- Note that the multipliers now are scale-dependent: the effects of final demand shocks will be different depending on the output level.

- Numerical simulation to explore the reaction of multipliers to changes in the output.

  - $x_{js} = x_{j2009} + N(0,0.1x_{j2009})$

  - 100 simulated outputs for each industry j.
\( \Delta \) Output of Electricity, Gas and Water Supply/\( \Delta \) final demand

\[ \frac{\text{Output of Electricity, Gas and Water Supply}}{\text{final demand}} \]

% of change from 2009 output in the industry
\( \Delta \) Output of Mining and Quarrying/\( \Delta \) final demand Elec., Gas and Water Sup.

% of change from 2009 output in Elec. industry

-30% -20% -10% 0% 10% 20% 30%
Table 2: GME estimates of the $\beta_j^r$ parameters

<table>
<thead>
<tr>
<th>Industry number</th>
<th>$\hat{\beta}_j^r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>i1</td>
<td>0.651**</td>
</tr>
<tr>
<td>i2</td>
<td>0.754**</td>
</tr>
<tr>
<td>i3</td>
<td>0.654**</td>
</tr>
<tr>
<td>i4</td>
<td>0.773**</td>
</tr>
<tr>
<td>i5</td>
<td>0.694**</td>
</tr>
<tr>
<td>i6</td>
<td>0.649**</td>
</tr>
<tr>
<td>i7</td>
<td>0.691**</td>
</tr>
<tr>
<td>i8</td>
<td>0.748**</td>
</tr>
<tr>
<td>i9</td>
<td>0.696**</td>
</tr>
<tr>
<td>i10</td>
<td>0.633**</td>
</tr>
<tr>
<td>i11</td>
<td>0.688**</td>
</tr>
<tr>
<td>i12</td>
<td>0.667**</td>
</tr>
<tr>
<td>i13</td>
<td>0.656**</td>
</tr>
<tr>
<td>i14</td>
<td>0.639**</td>
</tr>
<tr>
<td>i15</td>
<td>0.654**</td>
</tr>
<tr>
<td>i16</td>
<td>0.625**</td>
</tr>
</tbody>
</table>

** and * stands for estimates significantly different from 1 at 5% and 10% respectively
Δ labor in Construction/Δ final demand

% of change from 2009 output in the industry

Persons engaged

-30% -20% -10% 0% 10% 20% 30% 40%

9.4 9.6 9.8 10 10.2 10.4 10.6 10.8 11 11.2
5. Final remarks (and future research lines)

- We have the data and the estimation techniques to apply non linear analysis in our IO framework.

- Computationally more demanding but it would allow for more flexible analysis.

- Some points in the research agenda:
  
  - More factors in our equations.
  
  - Non linearities in the context of SUT’s.
THANK YOU

Esteban Fernández-Vázquez
University of Oviedo (Spain)
evazquez@uniovi.es