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

Fernandez-Vazquez; Esteban[a*]

[a] *RegioLab and Department of Applied Economics*
University of Oviedo, Spain
e-mail: evazquez@uniovi.es

*Corresponding author

Abstract

Non linear input-output (NIO) modelling, despite the relatively rich literature that developed its theoretical basis, has been only very modestly applied to empirical analysis. The main reason for this lack of empirical estimation of NIO models is that the number of parameters to estimate is much higher than the number of available data points. In order to solve this problem, calibration techniques are usually applied (as in the case of CGE estimation).

This paper proposes an alternative approach to estimate NIO models. Taking advantage of the proliferation of IO databases in the last few years; and by applying an estimation strategy that relies on entropy econometrics, the paper suggests estimating (instead of calibrating) the parameters that characterize a non-linear relation between inputs and output. This nonlinear model is characterized by having scale dependent input coefficients, instead of fixed ones. Several types of multiplier can be calculated from this nonlinear model, allowing for calculating confidence intervals of our results. The proposed technique is developed and then illustrated by means of an empirical application where the parameters that characterize a NIO model are estimated for the Spanish economy.

Keywords: Non linear input-output (NIO) modelling, scale dependent coefficients and multipliers, generalized maximum entropy

1. Introduction

Literature on non linear input output modeling is mainly theoretical. Researchers in this field have been concerned with the theoretical conditions required to prove the existence of solutions for the nonlinear IO models (see, for example, Lahiri, 1973; Lahiri and Pyatt, 1980; Chander, 1983; Fujimoto, 1986, or Dietzenbacher, 1994). On the other hand, examples of research where non linear models are empirically applied are particularly scarce, if we exclude the vast literature on computable general equilibrium models (CGE) where non linearities in the models are common. Some attempts of relaxing the assumptions of linear relations present in the standard Leontief IO models can be found in Heen (1992), West and Gamage (2001), Sun (2007) or Zhang (2008).

Perhaps the main motivation for the lack of empirical estimation of non linear IO models is that the number of parameters to estimate is much higher than the number of available data points. This restriction (solved in the field of CGE models by applying calibration techniques) has prevented IO researchers from conducting empirical applications of non linear IO models. The problem of limited information has been partially alleviated during the last years by the development of several projects that have released IO data at a world scale. Additionally, the use of alternative econometrics based on the statistical information theory has been growing in recent years as well. These estimators are characterized by performing comparatively better than traditional LS or ML estimators in contexts of small datasets.

This paper proposes the application of this type of estimators in order to estimate empirically the parameters characterizing a non-linear IO model. The paper is organized as follows. Section 2 presents the basic characteristics of a simple input-output model with scale-dependent coefficients. Section 3 provides a general description of the main characteristics of the estimator proposed for being used, and Section 4 provides detail of how it can be applied for estimating non-linear IO models. Section 5 presents an empirical application where such a non-linear IO model is estimated for the Spanish economy in 2009. Finally, Section 6 concludes the paper.

2. The nonlinear input-output model: scale dependent coefficients

The basic IO model assumes that input-output ratios are fixed; i.e., it is assumed that the output produced in one industry j (x_j) and the intermediate inputs from industry i that this industry j needs to produce its output (x_{ij}) are related by mean of a linear relation. These proportional relations are reflected in the technical coefficients $a_{ij} = x_{ij}/x_j$ by the expression:

$x_{ij} = (a_{ij})x_j$; being $0 \leq a_{ij} \leq 1$ (1)

In an economy with n industries the vector containing the outputs employed as intermediate input is then defined by $\boldsymbol{Ax}$, where $\boldsymbol{x}$ is the $(n \times 1)$ vector of output by industry and $\boldsymbol{A}$ stands for the $(n \times n)$ matrix with technical coefficients, which allows for defining the equation:

$\boldsymbol{x} = \boldsymbol{Ax} + \boldsymbol{f}$, (2)

being $\boldsymbol{f}$ the vector with the final demands by industry. Just reordering terms it is possible to derive the well known open Leontief model:

$\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A})^{-\mathbf{1}}\boldsymbol{f}$ (3)

where matrix $(\boldsymbol{I} - \boldsymbol{A})^{-\mathbf{1}}$ contains the output multipliers from changes in the final demand. One of the key assumptions in this model, and sometimes a source of criticism from non input-output practitioners, is that the coefficients a_{ij} are fixed and independent on the scale of production. This assumption can be relaxed by modifying equation (1) in the following terms:

$x_{ij}(x_j) = \alpha_{ij}x_j^{\beta_{ij}}$; being $\alpha_{ij}, \beta_{ij} \geq 0$ (4)

In other words, the linear relation between intermediate inputs and outputs is transformed in a scale-independent relation and equation (1) can be seen as a particular case of the non-linear equation (4) for $\beta_{ij} = 1$.[1]

This type of non-linear equations lead to a new type of scale-dependent technical coefficient defined as:

$a^*_{ij}(x_j) = x_{ij}/x_j = \alpha_{ij}x_j^{\beta_{ij}}/x_j = \alpha_{ij}x_j^{(\beta_{ij}-1)}$; (5)

In matrix notation this can be written in a new matrix of technical coefficients as $\boldsymbol{A}^*(\boldsymbol{x})$ that transforms the original linear Leontief model (3) into the non-linear equation:

$\boldsymbol{x} = \boldsymbol{A}^*(\boldsymbol{x})\boldsymbol{x} + \boldsymbol{f} = [\boldsymbol{I} - \boldsymbol{A}^*(\boldsymbol{x})]^{-\mathbf{1}}\boldsymbol{f}$ (6)

[1] Note that assuming that $\beta_{ij} = 1$ necessarily implies that $\alpha_{ij} = a_{ij}$.

A similar version of this non-linear IO model has been recently proposed by Roland-Holst and Sancho (2012). In their paper they show that under quite general conditions, equations such as the one in (6) have a unique non-negative solution $\boldsymbol{x}$ for any possible non-negative vector $\boldsymbol{f}$. These conditions can be summarized in the following set of assumptions:

i. $\boldsymbol{A}^*(\boldsymbol{x})\boldsymbol{x}$ is non-decreasing; implying that in order to produce more output, more intermediate inputs ($\boldsymbol{A}^*(\boldsymbol{x})\boldsymbol{x}$) are always required
ii. Continuity of $\boldsymbol{A}^*(\boldsymbol{x})\boldsymbol{x}$
iii. Expression (6) holds true for some pair $(\boldsymbol{f}, \boldsymbol{x})$.

Given the characteristics of (4) and assuming that parameters $\alpha_{ij}, \beta_{ij} \geq 0$, the matrix of scale-dependent coefficients $\boldsymbol{A}^*(\boldsymbol{x})$ in expression (6) holds assumptions (i) and (ii). Roland-Holst and Sancho (2012) show that assumption (iii) always holds if the model has been empirically implemented model by the base year solution.

Note that (6) states that the multipliers contained in the $[\boldsymbol{I} - \boldsymbol{A}^*(\boldsymbol{x})]^{-1}$ matrix are not scale-independent anymore: the effects derived by a change in the final demand will be different depending on the different levels of output by sector when this shock takes place. In other words, the multipliers that we estimate with this non-linear model will depend on the output levels. Parameters β_{ij}, in particular, play a key role here since their size determines the sign of the derivative:

$$\frac{da_{ij}^*(x_j)}{dx_j} = \alpha_{ij}(\beta_{ij} - 1)x_j^{(\beta_{ij}-2)} \tag{7}$$

From (4) and (7) is relatively easy to see that if $\beta_{ij} > 1$, this would be reflecting that the additional intermediate input x_{ij} necessary to produce additional output x_j is an increasing function of x_j (i.e., a production technology of x_j with decreasing returns to scale to intermediate input x_{ij}). This makes the scale-dependent technical coefficients $a_{ij}^*(x_j)$ increasing on the output levels, producing higher output multipliers at higher values of x_j. On the other hand, a parameter $\beta_{ij} < 1$ indicates that the additional quantities of intermediate input x_{ij} are decreasing on x_j (i.e., increasing returns to scale to intermediate input x_{ij}), making the $a_{ij}^*(x_j)$ decreasing on the output levels and results in lower output multipliers at higher values of x_j.

Frequently, IO models are used not only to calculate the output generated by a shock in the final demand, but also for deriving other type of multipliers (labor, CO_2 emissions, wages, etc.) generated by this type of shocks. Generally, this is done by an extension of (3) like:

$$\boldsymbol{l} = \hat{\boldsymbol{c}}\boldsymbol{x} = \hat{\boldsymbol{c}}(\boldsymbol{I} - \boldsymbol{A})^{-\mathbf{1}}\boldsymbol{f} = \boldsymbol{M}\boldsymbol{f} \tag{8}$$

Where $(n \times 1)$ vector $\boldsymbol{l}$ contains the values of the variable of interest and $\hat{\boldsymbol{c}}$ is a diagonal $(n \times n)$ matrix with elements $c_i = l_i/x_i$ in the main diagonal, representing the (labor, emissions, etc.) values of the variable of interest by unit of output. Matrix $\boldsymbol{M} = \hat{\boldsymbol{c}}(\boldsymbol{I} - \boldsymbol{A})^{-\mathbf{1}}$ contains the multipliers on variable $\boldsymbol{l}$ generated from changes in the final demand.

Similarly to the assumption made for the technical coefficients a_{ij}, conventional IO models assume that the c_i coefficients are independent on the scale of x_i. This linear relation between $\boldsymbol{l}$ and $\boldsymbol{x}$, can be transformed into something more flexible like:

$$l_j(x_j) = \alpha_j^l x_j^{\beta_j^l}; \tag{9}$$

Where normally α_j^l, β_j^l are assumed as non-negative. The new scale-dependent coefficient are now defined as:

$$c_j^*(x_j) = l_j/x_j = \alpha_j^l x_j^{\beta_j^l} \Big/ x_j = \alpha_j^l x_j^{(\beta_j^l - 1)}. \tag{10}$$

In matrix notation the non-linear model equivalent to (8) can be written now as:

$$\boldsymbol{l} = \hat{\boldsymbol{c}}^*(\boldsymbol{x})\boldsymbol{x} = \hat{\boldsymbol{c}}^*(\boldsymbol{x})\big(\boldsymbol{I} - \boldsymbol{A}^*(\boldsymbol{x})\big)^{-\mathbf{1}}\boldsymbol{f} = \boldsymbol{M}^*(\boldsymbol{x})\boldsymbol{f} \tag{11}$$

The new matrix of $\boldsymbol{l}$ multipliers $\boldsymbol{M}^*(\boldsymbol{x}) = \hat{\boldsymbol{c}}^*(\boldsymbol{x})\big(\boldsymbol{I} - \boldsymbol{A}^*(\boldsymbol{x})\big)^{-\mathbf{1}}$ is dependent on the scale of the output, being its values conditioned by the β_{ij} that affect matrix $\big(\boldsymbol{I} - \boldsymbol{A}^*(\boldsymbol{x})\big)^{-\mathbf{1}}$ and by parameters β_j^l affecting matrix $\hat{\boldsymbol{c}}^*(\boldsymbol{x})$, since:

$$\frac{dc_j^*(x_j)}{dx_j} = \alpha_j^l(\beta_j^l - 1)\, x_j^{(\beta_j^l - 2)} \tag{12}$$

We have the special case where $\beta_j^l = 1$, going back to the scale-independent coefficients $c_j^*(x_j) = c_j$ situation. The role of β_j^l parameters are similar to the original model with output multipliers, and (all other things being equal) a case where $\beta_j^l > 1$ produces coefficients $c_j^*(x_j)$

increasing on the output levels, resulting in higher output multipliers for higher values of x_j. The opposite situation happens when $\beta_j^l < 1$.

Given the importance of these parameters, a crucial point in such non-linear IO models is specifying possible values for them. For example, when constructing Computable General Equilibrium (CGE) models, non-linear relations between the variables are relatively common and calibration techniques are usually. Alternatively, this paper proposes an alternative approach and proposes an estimation strategy based on entropy econometrics, whose main characteristics are explained in the next section.

3. An overview of Entropy Econometrics

In spite of the proliferation of IO databases in recent years, we still do not have series of data large enough to apply traditional estimation techniques like Least Squares or Maximum Likelihood estimators, given the reduced numbers of degrees of freedom. Instead, we propose the application of Entropy Econometrics to estimate the parameters that characterize non linear IO models, given that these techniques have interesting properties when dealing ill-conditioned estimation problems (small samples). In Golan et al. (1996), Golan (2006) or Kapur and Kesavan (1992) extensive descriptions of the entropy estimation approach can be found.

Generally speaking, Entropy Econometrics techniques are used to recover unknown probability distributions of random variables that can take M different known values. The estimate $\tilde{\boldsymbol{p}}$ of the unknown probability distribution $\boldsymbol{p}$ must be as similar as possible to an appropriate *a priori* distribution $\boldsymbol{q}$, constrained by the observed data. Specifically, the Cross-Entropy (CE) procedure estimates $\tilde{\boldsymbol{p}}$ by minimizing the Kullback-Leibler divergence $D(\boldsymbol{p}\|\boldsymbol{q})$ (Kullback, 1959):

$$\underset{\boldsymbol{p}}{\text{Min}}\, D(\boldsymbol{p}\|\boldsymbol{q}) = \sum_{m=1}^{M} p_m \, ln\left(\frac{p_m}{q_m}\right) \tag{13}$$

The divergence $D(\boldsymbol{p}\|\boldsymbol{q})$ measures the dissimilarity of the distributions $\boldsymbol{p}$ and $\boldsymbol{q}$. This measure reaches its minimum (zero) when $\boldsymbol{p}$ and $\boldsymbol{q}$ are identical and this minimum is reached when no constrains are imposed. When the *a priori* distribution is not clearly defined because of lack of a prior information, the natural solution is assuming that in principle all the M values have the same probability, setting $\boldsymbol{q}$ as an uniform distribution. In this situation the CE procedure turns into the so-called Maximum Entropy methodology: a particular case of CE where the objective is to be as close as possible to the initial situation of maximum uncertainty. In such a case, minimizing equation (13) is equivalent to maximizing Shannon's entropy indicator defined by:

$$\underset{\boldsymbol{p}}{\text{Max}}\, H(\boldsymbol{p}) = -\sum_{m=1}^{M} p_m \, ln(p_m) \tag{14}$$

The underlying idea of the ME methodology can be applied for estimating the parameters of general linear models, which leads us to the so-called generalized Cross Entropy (GME). Let us suppose a variable y that depends on H explanatory variables x_h:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon} \tag{15}$$

Where $\boldsymbol{y}$ is a $(T \times 1)$ vector of observations for y, $\boldsymbol{X}$ is a $(T \times H)$ matrix of observations for the x_h variables, $\boldsymbol{\beta}$ is the $(H \times 1)$ vector of unknown parameters $\boldsymbol{\beta} = (\beta_1, \dots, \beta_H)$ to be estimated, and $\boldsymbol{\epsilon}$ is a $(T \times 1)$ vector with the random term of the linear model. Each β_h is assumed to be a discrete random variable. We assume that there is some information about its $M \geq 2$ possible realizations. This information is included for the estimation by means of a support vector $\boldsymbol{b}' = (b_1, \dots, b_M)$ with corresponding probabilities $\boldsymbol{p'}_h = (p_{h1}, \dots, p_{hM})$. The vector $\boldsymbol{b}$ is based on the researcher's a priori belief about the likely values of the parameter. For the sake of convenient exposition, it will be assumed that the *M* values are the same for every parameter, although this assumption can easily be relaxed. Now, vector $\boldsymbol{\beta}$ can be written as:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_H \end{bmatrix} = \boldsymbol{BP} = \begin{bmatrix} \boldsymbol{b'} & 0 & \cdots & 0 \\ 0 & \boldsymbol{b'} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{b'} \end{bmatrix} \begin{bmatrix} \boldsymbol{p_1} \\ \boldsymbol{p_2} \\ \vdots \\ \boldsymbol{p_H} \end{bmatrix} \tag{16}$$

Where $\boldsymbol{B}$ and $\boldsymbol{P}$ have dimensions $(H \times HM)$ and $(HM \times 1)$ respectively. Now, the value of each parameter β_h is given by the following expression:

$$\beta_h = \boldsymbol{b'p_h} = \sum_{m=1}^{M} b_m \, p_{hm}; \ \forall h = 1, \dots, H \tag{17}$$

For the random term, a similar approach is followed. Oppositely to other estimation techniques, GME does not require rigid assumptions about a specific probability distribution function of the stochastic component, but it still is necessary to make some assumptions. $\boldsymbol{\epsilon}$ is assumed to have mean $E[\boldsymbol{\epsilon}] = 0$ and a finite covariance matrix. Basically, we represent our uncertainty about the realizations of vector $\boldsymbol{\epsilon}$ treating each element ϵ_t as a discrete random variable with $J \geq 2$ possible outcomes contained in a convex set $\boldsymbol{v'} = \{v_1, \dots, v_J\}$, which for the sake of simplicity is assumed as common for all the ϵ_t. We also assume that these possible realizations are symmetric around zero ($-v_1 = v_J$). The traditional way of fixing the upper and lower limits of this set is to apply the three-sigma rule (see Pukelsheim, 1994). Under these conditions, vector $\boldsymbol{\epsilon}$ can be defined as:

$$\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_T \end{bmatrix} = \boldsymbol{VW} = \begin{bmatrix} \boldsymbol{v'} & 0 & \cdots & 0 \\ 0 & \boldsymbol{v'} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{v'} \end{bmatrix} \begin{bmatrix} \boldsymbol{w_1} \\ \boldsymbol{w_2} \\ \vdots \\ \boldsymbol{w_T} \end{bmatrix} \quad (18)$$

and the value of the random term for an observation t equals:

$$\epsilon_t = \boldsymbol{v'} \boldsymbol{w_t} = \sum_{j=1}^{J} v_j \, w_{tj}; \; \forall t = 1, \dots, T \quad (19)$$

Consequently, model (15) can be transformed into:

$$\boldsymbol{y} = \boldsymbol{XBp} + \boldsymbol{Vw} \quad (20)$$

So we need also to estimate the elements of matrix $\boldsymbol{W}$ (denoted by $\widetilde{w}_{tj}$) and the estimation problem for the general linear model (15) is transformed into the estimation of $H + T$ probability distributions. For this estimation the GME problem is written in the following terms:

$$\underset{\boldsymbol{P},\boldsymbol{W}}{\text{Max}}\, H(\boldsymbol{P}, \boldsymbol{W}) = -\sum_{h=1}^{H} \sum_{m=1}^{M} p_{hm} \, ln(p_{hm}) - \sum_{t=1}^{T} \sum_{j=1}^{J} w_{tj} \, ln(w_{tj}) \quad (21)$$

subject to:

$$y_t = \sum_{h=1}^{H} \sum_{m=1}^{M} b_m p_{hm} + \sum_{j=1}^{J} v_j \, w_{tj}; \;\; \forall t = 1, \dots, T \quad (22)$$

$$\sum_{m=1}^{M} p_{hm} = 1 \, ; \; \forall h = 1, \dots, H \quad (23)$$

$$\sum_{j=1}^{J} w_{tj} = 1 \, ; \; \forall t = 1, \dots, T \quad (24)$$

The restrictions in (22) ensure that the posterior probability distributions of the estimates and the errors are compatible with the observations. The equations in (23) and (24) are just normalization constraints.

4. Applying GME to estimate a non-linear IO model

The estimation strategy explained in the previous section can be applied to recover the parameters that determine the non linear relations present in equations like (6) or (11). In order to estimate the α_{ij} and β_{ij} parameters present in equations like (4), we need to collect information from a dataset (a time series or a cross-section) of IO tables and to assume that the production technology reflected in the parameters is constant across along all the observations (i.e., all the data points in the time series or the cross section of IO tables). Assuming that we have a dataset composed by T IO tables composed by the same n industries, for the basic IO model depicted in (6), n^2 equations like the following are estimated:

$$x_{ijt}(x_{jt}) = \alpha_{ij} x_{jt}^{\beta_{ij}}; t = 1, \ldots, T \tag{25}$$

When applying GME to estimate the type of equations (25), these equations are first linearized by means of a log transformation:

$$ln\,(x_{ijt}) = ln(\alpha_{ij}) + \beta_{ij}\, ln(x_{jt}) + \varepsilon_{ijt}; t = 1, \ldots, T \tag{26}$$

Where ε_t stands for the error term included in the equation, and then re-parameterized as follows:

$$ln\,(x_{ijt}) = ln(\sum_{m=1}^{M} a_m\, p_{am}) + \sum_{m=1}^{M} b_m\, p_{bm}\, ln(x_{jt}) + \sum_{j=1}^{J} v_j\, w_{tj}; t = 1, \ldots, T \tag{27}$$

Being a_m, b_m and v_j the points included in the supporting vectors for the α_{ij} and β_{ij} parameters and the error term ε_t respectively, with corresponding unknown probabilities p_{am}, p_{bm} and w_{tj}. Consequently, the following n^2 GME optimization programs need to be solved:

$$\text{Max}_{\boldsymbol{p_a},\boldsymbol{p_b},\boldsymbol{W}}\, H(\boldsymbol{p_a}, \boldsymbol{p_b}, \boldsymbol{W}) = -\sum_{m=1}^{M} p_{am}\, ln(p_{am}) - \sum_{m=1}^{M} p_{bm}\, ln(p_{bm}) - \sum_{t=1}^{T}\sum_{j=1}^{J} w_{tj}\, ln(w_{tj}) \tag{28}$$

subject to:

$$ln\,(x_{ijt}) = ln(\sum_{m=1}^{M} a_m\, p_{am}) + \sum_{m=1}^{M} b_m\, p_{bm}\, ln(x_{jt}) + \sum_{j=1}^{J} v_j\, w_{tj}; t = 1, \ldots, T \tag{29}$$

$$\sum_{m=1}^{M} p_{am} = \sum_{m=1}^{M} p_{bm} = 1\,;\, p_{am}, p_{bm} > 0 \tag{30}$$

$$\sum_{j=1}^{J} w_{tj} = 1\,;\ \forall t = 1, \ldots, T\,;\, w_{tj} > 0 \tag{31}$$

The estimates of the parameters will be given by the expressions:

$$\hat{\beta}_{ij} = \boldsymbol{b}'\boldsymbol{p_b} = \sum_{m=1}^{M} b_m \, p_{bm}; \; \forall i, j = 1, \dots, n \tag{32}$$

$$\hat{\alpha}_{ij} = \boldsymbol{a}'\boldsymbol{p_a} = \sum_{m=1}^{M} a_m \, p_{am}; \; \forall i, j = 1, \dots, n \tag{33}$$

Similarly, if we extend our IO model to estimate multipliers of variable $\boldsymbol{l}$, we also need to estimate n equations like:

$$l_{jt}(x_{jt}) = \alpha_j^l x_{jt}^{\beta_j^l}; \tag{34}$$

Which, after the convenient linearization and the inclusion of an error term are written as:

$$ln\,(l_{jt} =)\, ln(\alpha_j^l) + \beta_j^l \, ln(x_{jt}) + \epsilon_{jt}; \, t = 1, \dots, T \tag{35}$$

After the same type of re-parameterization as the one applied before, the n GME programs are:

$$\text{Max}_{\boldsymbol{p_a^l}, \boldsymbol{p_a^l}, \boldsymbol{U}} H(\boldsymbol{p_a^l}, \boldsymbol{p_a^l}, \boldsymbol{U}) = -\sum_{m=1}^{M} p_{am}^l \, ln(p_{am}^l) - \sum_{m=1}^{M} p_{bm}^l \, ln(p_{bm}^l) - \sum_{t=1}^{T} \sum_{j=1}^{J} u_{tj} \, ln(u_{tj}) \tag{36}$$

subject to:

$$ln\,(l_{jt}) = ln\left(\sum_{m=1}^{M} a_m^l \, p_{am}\right) + \sum_{m=1}^{M} b_m^l \, p_{bm} \, ln(x_{jt}) + \sum_{j=1}^{J} v_j^l \, u_{tj}; \, t = 1, \dots, T \tag{37}$$

$$\sum_{m=1}^{M} p_{am}^l = \sum_{m=1}^{M} p_{bm}^l = 1 \, ; \, p_{am}^l, p_{bm}^l > 0 \tag{38}$$

$$\sum_{j=1}^{J} u_{tj} = 1 \, ; \, \forall t = 1, \dots, T \, ; \, u_{tj} > 0 \tag{39}$$

Now a_m^l, b_m^l and v_j^l stand for the points included in the supporting vectors for the α_j^l and β_j^l parameters and the error term ϵ_t respectively. Their corresponding probabilities to be estimated are p_{am}^l, p_{bm}^l and u_{tj}; which once the GME programs is solved, allow for getting estimates of the parameters like:

$$\hat{\beta}_j^l = \boldsymbol{b_l}'\boldsymbol{p_{bl}} = \sum_{m=1}^{M} b_m^l \, p_{bm}^l; \; \forall j = 1, \dots, n \tag{40}$$

$$\hat{\alpha}_j^l = \boldsymbol{a_l}'\boldsymbol{p_{al}} = \sum_{m=1}^{M} a_m^l \, p_{am}^l; \; \forall j = 1, \dots, n \tag{41}$$

5. An illustration: output and labor multipliers in the Spanish economy with a non-linear IO model for 2009

The empirical estimation of these non-linear IO models will be illustrated in this section by estimating the impact in terms of creation of jobs of changes in the final demand of the Spanish economy. The IO table taken as reference for this empirical exercise is the industry-by-industry IO table published in the WIOD project database corresponding to 2009.[2] This particular database has been taken for this analysis for several reasons. One is that it provides more recent information than the Spanish Statistical Institute (INE), which only publishes harmonized industry-by-industry tables until 2005. Second, the WIOD database allows for using a series of IO tables for Spain from 1995 until 2009 with an homogenous sectoral classification along time and with information about employment (among other economic and environmental indicators) following the same classification.

Even when the sectoral classification in these IO tables allows for identifying 34 different industries, a sectoral aggregation into a new classification with 16 economic branches has been considered instead, in order to reduce the computational burden and to ease the presentation of results. The specific sectoral classification used can be found in the Appendix.

First a model like (6) has been considered. The information required to estimate this model are the interindustry transactions (in current million \$) x_{ij} together with the total industry inputs x_j in the series of IO tables from 1995 to 2009. In particular, the 16^2 equations to be estimated are:

$$ln\ (x_{ijt}) = ln\left(\alpha_{ij}\right) + \beta_{ij}\ ln\left(x_{jt}\right) + \varepsilon_{ijt};\ t = 1995,\ldots,2009 \qquad (42)$$

But they are modified as follows in order to prevent possible spurious regression in the time series of observations:

$$\Delta ln\ (x_{ijt}) = \Delta\beta_{ij}\ ln\left(x_{jt}\right) + \varepsilon_{ijt};\ t = 1996,\ldots,2009 \qquad (43)$$

The GME program used in this estimation is the same type as the one described in equations (28) to (31) with the particularity that the GME program does not estimates the α_{ij} directly, but they are specified as $ln\left(\alpha_{ij}\right) = \overline{ln\ (x_{ij})} - \beta_{ij}\ \overline{ln\left(x_j\right)}$.[3] This implies that setting supporting vectors for the possible values for this parameter is not necessary. We still need, however,

[2] See www.wiod.org for details.
[3] This expression is included as a constrain in the GME program directly.

specifying a supporting vector for the parameters β_{ij} but this specification can be done in a relatively "natural" way. More specifically, a support vector with $M = 3$ points like $\boldsymbol{b}' = (b_1, b_2, b_3)$ has been considered. This allows for setting the central point b_2 as the "uninformative" estimate (i.e., the estimate that we would expect if no information was available) and then setting the bounds symmetrically around it. In a situation with no additional information, one would expect that this parameter was equal to 1, since there would not be any evidence to depart from the simple linear IO model.[4] On this basis, the central value has been specified as $b_2 = 1$. The lower limit b_1 can be set easily as well, since we do not allow for negative values of the parameters and therefore $b_1 = 0$. Consequently, since the values are set symmetrically around the central one, this implies that $b_3 = 2$. For the supporting vector of the error terms, the traditional three-sigma rule has been applied.

Additionally, given that a desirable objective is that the estimated matrix of scale-dependent technical coefficients was composed by elements characterized by being positive and with column sums smaller than one, these two additional constrains have been included in the GME program:

$$\alpha_{ij} = \sum_{m=1}^{M} a_m \, p_{am} > 0; \; \forall i, j = 1, \dots, n \tag{44}$$

$$\sum_{i=1}^{n} \alpha_{ij} x_j^{(\beta_{ij}-1)} < 1; \; \forall j = 1, \dots, n \tag{45}$$

By solving such a GME program, the scale-dependent coefficients $a_{ij}^*(x_j)$ are estimated at the output industry levels corresponding to 2009.

In a similar fashion, a non linear IO model like (11) for quantifying the effect of final demand shocks in the variations of jobs by industry has been estimated as well. The data on which this estimation is based on are the number of persons engaged by industry, obtained from the WIOD database for Spain 1995-2009 as well. Now, a GME program like the one depicted in equations (36) to (39) is implemented, including as constrain the information that relates output levels with the number of jobs by industry in our sample by the expression:

$$\Delta ln\,(l_{jt}) = \Delta\beta_j^l \, ln(x_{jt}) + \epsilon_{jt}; \, t = 1996, \dots, 2009 \tag{46}$$

[4] Note that this implies that the estimation of the non-linear IO model without the sample information contained in equations like (29) will get as solution the traditional linear IO model.

The same logic as in the GME program for estimating the previous non linear IO model applies here as well. A support vector with $M = 3$ points has been considered again, setting the central value as $b_2 = 1$ and the lower and upper limits as $b_1 = 0$ and $b_3 = 2$ respectively. The traditional three-sigma rule has been applied for the supporting vector of the error terms.

Note that our estimates of these coefficients are by definition exactly equal to the traditional technical and labor coefficients for this "base" year if the estimates of the errors are considered as well:

$$a^*_{ij2009}(x_j) = \hat{\alpha}_{ij} x_{j2009}^{(\hat{\beta}_{ij}-1)} exp(\hat{\varepsilon}_{ij2009}), \tag{47}$$

$$c^*_{j2009}(x_{j2009}) = \hat{\alpha}^l_j x_{j2009}^{(\hat{\beta}^l_j-1)} exp\ (\hat{\epsilon}_{j2009}); \tag{48}$$

but it allows for estimating different coefficients (and therefore different output and job multipliers) if the industry outputs are assumed to change from the 2009 levels. Denoting by x_{js} the simulated output levels, these estimates will be defined by the expressions:

$$a^*_{ijs}(x_{js}) = \hat{\alpha}_{ij} x_{js}^{(\hat{\beta}_{ij}-1)} \tag{49}$$

$$c^*_{js}(x_{js}) = \hat{\alpha}^l_j x_{js}^{(\hat{\beta}^l_j-1)} \tag{50}$$

Following these ideas, the non linear matrices of output and jobs multipliers, $\left(I - \boldsymbol{A}^*(\boldsymbol{x})\right)^{-\mathbf{1}}$ and $\boldsymbol{M}^*(\boldsymbol{x})$ respectively, have been estimated for the Spanish economy from the mentioned series of IO tables from 1995 to 2009.

<<Insert Tables 1 and 2 about here>>

Tables 1 and 2 show the GME estimates of the β_{ij} and β^l_j parameters respectively. Focusing in Table 1, there are not large differences with respects of the linear model ($\beta_{ij} = 1$), but one can still detect some patterns.[5] For example, the sector composed by the Coke, Refined Petroleum and Nuclear Fuel industries (column i6) seems to present systematically estimates smaller than one, indicating that increases in the output of this sector derived from a positive shock in its final demand will result in smaller effects generated on the output of other sectors if the output when the output of this industry is relatively high. A similar situation, but to a lesser

[5] The highest 10% of estimates in Table 1 are written in bolds, whereas the smallest 10% are underlined and written in italics

extent can be observed for the branches of Electricity, Gas and Water Supply (i10) and FIRE services and other business activities (i15).

Oppositely, it is possible to identify some industries where the additional intermediate inputs required to produce additional output is an increasing function of x_j. This can be seen, for example, for the sector composed by the Hotel and Restaurant industries (column i13), where the β_{ij} estimates corresponding to the intermediate inputs from Transport, post and telecommunications (i14) or FIRE services and other business activities (i15) are among the highest ones. Something similar happens with the Sale, maintenance and trade industries (column i12), with estimates of this parameter larger than one for the intermediate output from Basic Metals and Fabricated Metal (i8), Machinery, equipment and n.e.c. manufacturing (i9), Electricity, Gas and Water Supply (i10) and Construction (i11). In all these cases the scale-dependent technical coefficients $a_{ij}^{*}(x_j)$ are estimated to be increasing on the output levels, producing higher output multipliers at higher values of the industry outputs.

Table 2 reports the estimates for the β_j^l parameters. Again the results are not very different form the unit values, but they allow for identifying two groups of industries. Only Agriculture, Hunting, Forestry and Fishing (i1), Wood, pulp and paper (i5), Chemicals, rubber, plastics and non-metallic mineral (i7), Construction (i11) and, specially, Textiles, Leather and Footwear products (i4) present estimates larger than 1, indicating that only for these cases labor coefficients $c_j^{*}(x_j)$ increasing on the output levels are estimated. Note that this implies that for these industries one should expect higher output multipliers for higher values of their outputs, and the opposite case happens for the remaining 12 industries.

In order to illustrate the usefulness of such non-linear IO models, a small numerical experiment has been carried out. This simulation will generate 100 different output levels (x_{js}) by industry departing from the original values at 2009. These simulated outputs are generated as $x_{js} = x_{j2009} + N(0{,}0.1x_{j2009})$. Then, the non-linear intermediate inputs and labor coefficients, $a_{ijs}^{*}(x_{js})$ and $c_{js}^{*}(x_{js})$ respectively, are calculated based on equations (49) and (50). This will allow estimating a matrix of labor multipliers defined as $\boldsymbol{M}^{*}(\boldsymbol{x}) = \hat{\boldsymbol{c}}^{*}(\boldsymbol{x})\left(\boldsymbol{I} - \boldsymbol{A}^{*}(\boldsymbol{x})\right)^{-\mathbf{1}}$ that will be different for each simulated output. The mean and the standard deviation of each element $m_{ijs}^{*}(x_{js})$ on this matrix are computed along the simulations, allowing for computing the bounds of "confidence intervals" for the labor multipliers defined as the mean $\pm$ three times the standard deviation. Table 3 shows the difference (in %) between the upper and lower bounds for each cell of matrix $\boldsymbol{M}^{*}(\boldsymbol{x})$.

<<Insert Table 3 about here>>

Results in Table 3 seem to suggest that the variability in the estimates of the labor multipliers can be substantial depending on the output level for which these multipliers are calculated. [6] Not surprisingly, the sector of Coke, Refined Petroleum and Nuclear Fuel industries (column i6), where the estimates of the β_{ij} parameter where remarkably different (smaller) from 1, and the sector of Textiles, Leather and Footwear products (row i4), where the β_j^l parameter had the highest estimate are those presenting the largest differences between the bounds.

Just to have a clue about what differences in the calculation of the labor multipliers can represent in jobs, Figures 1 and 2 plot the estimated $m_{jjs}^{*}(x_{js})$ elements (i.e., the additional jobs in industry j generated by an increase in the final demand of this industry j) for the sectors of Construction (i11) and FIRE services and other business activities (i15). Specifically, these figures show how many jobs (in thousands of persons engaged) are generated by all the multiplier effects derived of 1 billion of US$ increase on their final demand.

<<Insert Figures 1 and 2 about here>>

The vertical axis in both graphs represents the total additional jobs generated, whereas the horizontal axis contains the output levels simulated expressed as a % difference with respect to the 2009 levels. The grey points show the job multipliers estimated for each level of output, whereas the horizontal black line marks the reference of the estimate under a traditional linear IO model like (11). These two industries taken for the example illustrate how the use of such non-linear IO modelling can add some additional information to the standard derivation of IO multipliers: it allows for establishing a range of plausible values for the estimated multipliers (the impact on the number of additional jobs created ranges approximately between 14,600 and 15,600 in the Construction industry or between 7,800 and 8,200 in the case of FIRE services and other business activities).

[6] The highest 10% of estimates in Table 3 are written in bolds, whereas the smallest 10% are underlined and written in italics

6. An illustration: output and labor multipliers in the Spanish economy with a non-linear IO model for 2009

This paper basically presents a proposed technique to estimate empirically the parameters characterizing a non-linear IO model with scale-dependent coefficients. The idea of the non-linear IO model itself is not new itself, but it has been deserved some attention in the IO literature. However, this attention has been normally focused on theoretical characteristics of the model (see Chander, 1983; or Fujimoto, 1986; among others), but not in its possible empirical implementation.

The lack of datasets large enough for applying traditional econometrics has prevented IO researchers from conducting empirical applications of such models, with the exception of CGE practitioners who base their computations on an IO model in the core of their models and the calibration (not estimation) of the necessary parameters. However, during the last years several projects for developing global IO databases have flourished (like WIOD, Exiopol or Eora), which can be helpful for expanding the sample sizes with the required data to conduct empirical estimations of non linear IO models. Moreover, alternative estimation techniques can be safely applied even in situations with relative small datasets. Entropy Econometrics, in general, and the technique known as Generalized Maximum Entropy (GME), in particular, can be considered for its potential application in this field.

In this paper it is shown how the GME estimator can be applied for estimating IO models characterized for having scale dependent coefficients. Its flexibility allows for including all the data available in one sample, keeping at the minimum the distributional assumptions about the error, but including at the same time all the constrains necessary to guarantee that the estimated model holds a series of properties. Its performance is illustrated by estimating a non linear IO model for quantifying the labor multipliers by industry in the Spanish economy, taking advantage of the homogenous series of IO tables compiled in the WIOD project. The results obtained by means of a small numerical experiment show the potential of such type of modeling, since it allows for calculating “confidence intervals” to the estimates of the multipliers, which can contribute to enrich the analysis performed by a traditional IO model with fixed coefficients.

Even when the results presented in this paper are encouraging, this line of research still has several topics to study in the future. One important issue that should be considered is the robustness of the estimates obtained by the proposed estimator to changes in some of the assumptions. When GME estimators are applied to datasets characterized by small sample sizes, sensitivity analysis are useful to quantify the effect of the a priori information imposed by the researcher on the outcomes of the estimation: if the results are highly sensitive to the non-sample information included in the support vectors, this is a signal that the GME estimates are not robust. This type of questions, together with additional empirical applications not based only on a time series for a single country, but considering a panel of homogenous IO tables should be in the center of the research agenda on this field.

References

Chander, P. (1983). The Nonlinear Input-Output Model, Journal of Economic Theory, 30, pp. 219-229.

Dietzenbacher, E. (1994). "The non-linear Perron-Frobenius theorem: Perturbations and aggregation," Journal of Mathematical Economics, vol. 23(1), pp. 21-31.

Fujimoto, T. (1986). “Non-linear Leontief Models in Abstract Spaces”, Journal of Mathematical Economics, 15, pp.151-156.

Golan, A. (2006): Information and entropy econometrics. A review and synthesis, Foundations and Trends in Econometrics, 2, pp. 1-145.

Golan, A., Judge, G. and D. Miller, (1996): Maximum Entropy Econometrics: Robust Estimation with Limited Data, Nueva York, John Wiley & Sons.

Heen, K. (1992): Impact Analysis with Variable Input-Output Coefficients, Economic Systems Research, 4(2) pp. 115-125.

Kapur Jn, Kesavan Hk, (1992): Entropy Optimization Principles with Applications. Academic Press. New York.

Lahiri, F. (1973) Input-output analysis with scale-dependent coefficients, Econometrica, 44 (2), pp. 947-961.

Roland-Holst, D. W., and Sancho, F. (2012): An operational, nonlinear input-output system. Paper presented the 20th International Input-Output Conference. Bratislava, Slovakia, June 24-29.

Sun Y. Y. (2007): Adjusting Input–Output models for capacity utilization in service industries. Tourism Management, 28, pp. 1507-1517.

West, G., and Gamage, A. (2001): Macro effects of tourism in Victoria, Australia: A nonlinear input–output approach. Journal of Travel Research, 40(1), pp. 101–109.

Zhang J. S. (2008): A Multi-sector Nonlinear Dynamic Input–Output Model with Human Capital, Economic Systems Research, 20(2) pp. 223-237.

Table 1: GME estimates of the $\boldsymbol{\beta_{ij}}$ parameters

$\hat{\beta}_{ij}$	**i1**	**i2**	**i3**	**i4**	**i5**	**i6**	**i7**	**i8**	**i9**	**i10**	**i11**	**i12**	**i13**	**i14**	**i15**	**i16**
i1	1.048	0.933	0.986	0.937	0.991	*0.589*	**1.057**	0.872	0.964	*0.771*	*0.806*	0.922	0.897	0.959	1.048	0.933
i2	1.026	**1.083**	1.052	**1.073**	1.055	0.957	**1.097**	1.024	**1.060**	*0.718*	0.949	0.983	1.004	1.002	1.026	**1.083**
i3	0.956	0.917	0.967	0.846	0.937	*0.691*	0.932	0.874	*0.769*	0.935	*0.824*	0.923	1.008	0.893	0.956	0.917
i4	1.004	0.991	1.017	1.039	1.039	*0.615*	0.993	1.017	0.973	0.944	1.014	0.983	0.973	1.010	1.004	0.991
i5	1.041	0.919	0.917	1.052	0.978	*0.405*	0.984	0.966	0.962	0.956	0.903	0.981	0.923	*0.843*	1.041	0.919
i6	0.859	0.943	1.032	**1.065**	1.014	**1.084**	0.938	0.981	0.959	0.935	*0.804*	0.963	0.933	0.932	0.859	0.943
i7	0.935	0.870	1.020	0.939	0.956	*0.560*	0.980	0.903	0.936	0.889	0.913	0.995	0.923	0.922	0.935	0.870
i8	1.019	1.002	0.975	**1.067**	1.023	*0.452*	**1.064**	1.006	1.001	0.965	0.962	1.052	0.952	0.984	1.019	1.002
i9	1.002	1.049	**1.070**	**1.064**	1.008	*0.623*	1.038	1.021	**1.056**	0.943	0.984	1.051	1.000	1.037	1.002	1.049
i10	0.992	0.993	**1.072**	1.029	1.021	*0.552*	1.053	1.010	0.900	**1.069**	0.928	1.044	1.043	0.981	0.992	0.993
i11	1.002	0.996	0.986	1.015	1.016	*0.562*	1.040	0.956	0.936	1.017	**1.074**	1.039	0.984	**1.078**	1.002	0.996
i12	0.999	0.933	1.010	0.926	1.038	0.857	0.995	0.942	0.913	0.915	0.911	0.986	1.034	1.034	0.999	0.933
i13	1.046	0.921	1.044	1.025	0.988	*0.488*	0.987	0.884	0.931	0.982	0.982	0.958	1.002	0.987	1.046	0.921
i14	1.055	1.003	1.052	0.877	0.982	*0.511*	0.982	0.933	*0.840*	**1.067**	0.951	1.034	**1.074**	1.045	1.055	1.003
i15	1.019	1.021	0.957	0.922	0.968	*0.378*	1.001	0.906	0.891	1.024	0.980	**1.056**	**1.063**	1.042	1.019	1.021
i16	0.993	0.916	1.048	0.953	1.036	*0.475*	1.001	0.910	*0.841*	1.026	0.990	0.995	**1.061**	**1.073**	0.993	0.916

Table 2: GME estimates of the $\boldsymbol{\beta_j^l}$ parameters

Industry number	$\widehat{\boldsymbol{\beta}}_j^l$
i1	1.135
i2	0.857
i3	0.970
i4	1.470
i5	1.115
i6	0.982
i7	1.037
i8	0.979
i9	0.948
i10	0.864
i11	1.100
i12	0.868
i13	0.948
i14	0.781
i15	0.874
i16	0.834

Table 3: % differences between the upper and lower bounds of the estimates of $m^{*}_{ijs}(x_{js})$

	i1	**i2**	**i3**	**i4**	**i5**	**i6**	**i7**	**i8**	**i9**	**i10**	**i11**	**i12**	**i13**	**i14**	**i15**	**i16**
i1	7.38%	7.41%	7.65%	9.77%	7.38%	25.00%	7.69%	8.93%	8.86%	8.33%	10.10%	8.60%	7.93%	7.46%	7.69%	9.76%
i2	15.38%	7.91%	10.00%	11.11%	5.26%	8.33%	9.88%	7.79%	12.00%	18.13%	8.57%	8.33%	10.00%	8.33%	11.11%	10.00%
i3	2.99%	5.56%	2.23%	8.82%	2.27%	16.67%	5.13%	5.00%	9.09%	*0.00%*	5.00%	3.57%	2.15%	*0.00%*	7.14%	2.27%
i4	**36.84%**	**38.89%**	**34.48%**	**37.15%**	**34.48%**	**66.67%**	**35.71%**	**36.00%**	**37.70%**	30.00%	**35.00%**	**37.50%**	**38.46%**	**42.11%**	**33.33%**	**36.00%**
i5	6.49%	7.50%	8.11%	6.90%	6.72%	**40.00%**	7.03%	6.72%	7.35%	7.14%	7.37%	6.25%	7.77%	7.14%	9.48%	7.91%
i6	*0.00%*	9.09%	*0.00%*	*0.00%*	*0.00%*	*0.80%*	11.11%	*0.00%*	**50.00%**	8.33%	*0.00%*	*0.00%*	**50.00%**	8.33%	*0.00%*	*0.00%*
i7	3.72%	5.99%	*2.10%*	4.03%	3.49%	30.30%	*1.91%*	5.62%	3.78%	4.81%	4.40%	*1.79%*	2.13%	3.17%	3.19%	2.44%
i8	*1.60%*	*1.34%*	*1.80%*	2.13%	*1.67%*	**35.14%**	3.08%	*1.39%*	*1.49%*	2.20%	*1.62%*	2.17%	*1.22%*	2.44%	*1.41%*	*1.56%*
i9	2.80%	2.77%	3.08%	3.02%	2.95%	25.93%	3.24%	3.16%	2.74%	3.77%	2.82%	3.38%	3.23%	3.74%	3.26%	2.84%
i10	6.00%	7.38%	7.14%	7.94%	7.32%	**33.33%**	7.77%	6.74%	7.41%	7.16%	8.11%	6.56%	7.50%	6.38%	7.69%	6.82%
i11	8.61%	8.75%	8.80%	8.93%	8.74%	**32.14%**	8.76%	9.76%	10.04%	8.52%	8.78%	9.21%	9.04%	9.04%	8.69%	8.80%
i12	7.66%	8.19%	7.66%	8.71%	8.03%	15.35%	7.69%	8.96%	9.14%	8.48%	8.60%	7.63%	7.70%	8.46%	8.01%	7.49%
i13	2.86%	4.29%	2.74%	5.08%	3.03%	**33.33%**	3.03%	6.35%	5.17%	2.17%	2.82%	2.70%	2.85%	2.97%	4.40%	3.45%
i14	13.62%	13.37%	13.58%	16.34%	13.60%	**38.40%**	13.41%	13.42%	14.82%	13.40%	13.23%	13.99%	13.89%	13.38%	13.83%	13.86%
i15	7.58%	7.57%	7.74%	8.57%	7.64%	**40.38%**	7.45%	9.17%	9.14%	7.59%	7.46%	7.62%	7.68%	7.85%	7.50%	7.95%
i16	10.42%	10.84%	10.57%	11.02%	10.52%	**37.74%**	10.22%	12.70%	12.39%	10.53%	10.48%	10.35%	10.26%	11.14%	10.30%	10.23%

Figure 1: estimates of labor multipliers for Construction (i11) in the simulation

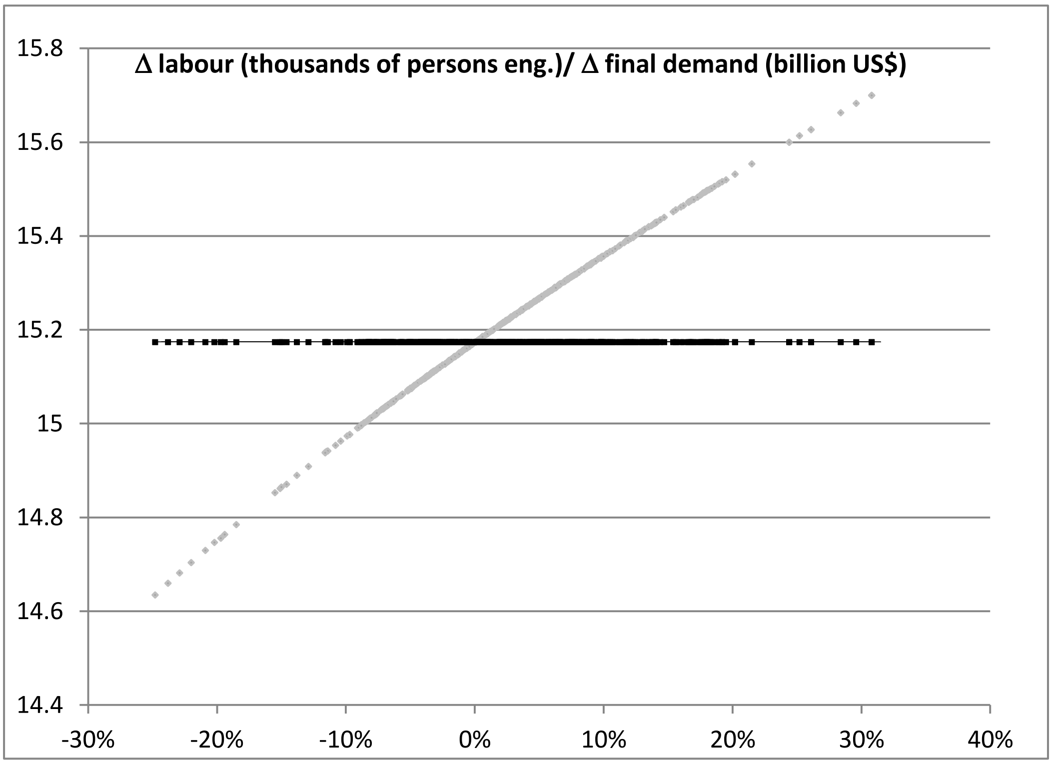

Figure 2: estimates of labor multipliers for FIRE services and other business activities (i15) in the simulation

Appendix: industry classification

Industry number	Industry description
i1	Agriculture, Hunting, Forestry and Fishing
i2	Mining and Quarrying
i3	Food, Beverages and Tobacco
i4	Textiles, Leather and Footwear products
i5	Wood, pulp and paper
i6	Coke, Refined Petroleum and Nuclear Fuel
i7	Chemicals, rubber, plastics and non-metallic mineral
i8	Basic Metals and Fabricated Metal
i9	Machinery, equipment and n.e.c. manufacturing
i10	Electricity, Gas and Water Supply
i11	Construction
i12	Sale, maintenance and trade
i13	Hotels and Restaurants
i14	Transport, post and telecommunications
i15	FIRE services and other business activities
i16	Other services