Using Additional Information in Structural Decomposition Analysis: The Path Based Approach

Esteban Fernández Vázquez[1], Bart Los[2,*] and Carmen Ramos Carvajal[1]

[1]University of Oviedo, Department of Applied Economics, Faculty of Economics, Campus del Cristo, Oviedo, 33006, Spain. e-mail: evazquez@uniovi.es; cramos@uniovi.es

[2]University of Groningen, Groningen Growth and Development Center and SOM Research School, P.O. Box 800, NL-9700 AV Groningen, The Netherlands. e-mail: b.los@eco.rug.nl

*Corresponding author.

ABSTRACT

Structural decomposition analysis (SDA) is a well-known methodology to assess the relative importance of effects that together constitute the actual change in a variable of interest. A widely recognized problem of SDA is that the results often depend strongly on the specific decomposition formula chosen, while numerous formulae are equivalent from a theoretical point of view. This "non-uniqueness" problem is often solved rather pragmatically, by reporting an average of (a subset of) all possible formulae. In this paper, we propose an approach that uses maximum entropy econometrics techniques to incorporate additional information to choose a specific decomposition formula. We illustrate the method empirically by investigating the sources of change in sectoral real labor costs in Spain, 1980-1994.

Keywords: Structural decomposition analysis, maximum entropy econometrics, labor costs, Spain.

Acknowledgements: This paper is based on Chapters 2 and 3 of the first author's Ph.D. thesis (Fernández, 2004). The authors would like to thank Erik Dietzenbacher for his comments on parts of the thesis that are relevant for this paper.

1. Introduction

A most basic input-output model allows to predict (*ex ante*) to what extent gross output levels are affected by changes in the final demand vector or the input-output matrix. In fact, the first input-output tables were constructed to make such impact analyses possible. The same basic model enables us to assess *ex post* the actual contributions of changes in these underlying factors to observed changes in gross output levels. The latter type of analysis has been coined structural decomposition analysis (SDA). An extensive overview of the methodology and its relatively early applications was provided by Rose & Casler (1996). More recent applications of different sorts include De Haan (2001), Hoekstra & Van den Bergh (2002) and Dietzenbacher *et al.* (2000, 2004). As was shown by Dietzenbacher & Los (1998, 2000), typical SDA results should be taken with care, because several methodological problems pertain to the techniques employed. This paper aims at providing a new perspective on the so-called "non-uniqueness" problem. Dietzenbacher & Los (1998) showed that this problem, which will be outlined below, is not only of theoretical interest. It is also an issue that is important in an empirical sense. Relative contributions of distinct sources of change (such as final demand change and changes in input coefficients, in the above-mentioned simple example) were shown to depend considerably on the specific decomposition form chosen.

If one decomposition form would be preferable over the others, the problem would not be relevant. Dietzenbacher & Los (1998) argue that the number of theoretically equivalent forms amounts to $n!$, in which n represents the number of distinct sources of change. Their admittedly pragmatic solution is to present averages of results obtained for all decomposition forms. They also show that it is often defendable from an empirical point of view to report averages over a well-defined small subset of forms. Nevertheless, an explicit choice for a specific decomposition form, or in other words, a specific attribution of interaction terms to the distinct sources of change is not made.

This paper proposes a different perspective. We argue that any available additional information for periods inbetween the initial and final time period considered can be used to divide the interaction terms in a way that fits the data better than implied by simply taking averages. The additional data are used in a Maximum Entropy (ME) estimation procedure to arrive at parameter estimates that together specify a unique division of the interaction terms.[1]

The paper is organized as follows. In Section 2, we present the "non-uniqueness" problem in SDA in formal terms. An illustration of a simple decomposition analysis with two determinants shows that the two single decomposition forms are special cases of a much broader class of divisions of the interaction term introduced in Section 3. So is the average, but this only holds for

[1] Maximum Entropy econometrics and strongly related Cross Entropy methods have been used in an intersectoral setting before. See, for example, Golan *et al.* (1994) and Robinson *et al.* (2001) for methods to estimate missing data in input-output tables and social accounting matrices.

the case with two determinants. We will show that the class can be represented by two simple equations. Specific divisions of the interaction term are characterized by two parameters. We will also show that the appoach can relatively easily be generalized to cases with more than two determinants. In Section 4, the principles of ME estimation are highlighted, and we show how ME estimation techniques can be used to estimate the parameter of interest in solving the "non-uniqueness" problem in SDA. Section 5 is devoted to a discussion of two types of additional information that can be used to implement the ME approach. In Section 6 we present an empirical illustration of the approach. We will study changes in sector-level labor costs in Spanish sectors between 1980 and 1994. Our aim is to assess the importance of sector-specific changes in the labor costs per unit of output on the one hand, and structural effects as a consequence of changes in the matrix of input coefficients and the vector of final demands on the other. We show that the most likely division of the interaction term within the class of divisions considered often deviates substantially from the average of all forms analyzed by Dietzenbacher & Los (1998), among others. Consequently, the shares of the total change attributed to specific determinants will be different as well. Section 7 concludes.

2. The Non-Uniqueness Problem

In input-output analysis, the most basic equation is $\mathbf{q}=\mathbf{Lf}$. Sectoral gross output levels $\mathbf{q}$ are expressed as the product of the Leontief inverse matrix $\mathbf{L}$ and the vector of sectoral final demand levels $\mathbf{f}$. Differences in gross output levels (be it over time or over countries or regions) can thus be due to differences in the values contained in the Leontief inverse and/or to differences in the final demand levels. The basic aim of SDA is to quantify the part of the differences in $\mathbf{q}$ that can be attributed to differences in $\mathbf{L}$ and the part caused by differences in $\mathbf{f}$. Because this problem is very much alike problems encountered in other subdisciplins in economics, we prefer to explain the more traditional approaches and our new approach in terms of a more general notation.[2]

Assume that the value of an endogenous variable z is given as the product of a set of n exogenous variables (or, determinants) x_1, x_2, ..., x_n.[3] That is:

$$z = x_1 x_2 ... x_n \tag{1}$$

[2] Actually, Fernández (2004, Chapter 4) offers an application of our methodology to a shift-share analysis of employment growth in Spanish regions.

[3] The endogenous and exogenous variables can be represented by scalars, vectors and/or matrices. Throughout the paper, we adopt the convention that scalars are represented by italic lowercase symbols, (column) vectors by lowercase bold symbols and matrices by bold capitals. Primes denote transposition and hats indicate diagonal matrices.

A fundamental assumption is that the exogenous variables can be assumed to be independent, not only in a mathematical sense (see Dietzenbacher and Los, 2000, for an account of problems related to mathematical dependency of determinants) but also from an economic-theoretical viewpoint. That is, we assume that each determinant could change without an necessarily accompanying change in the values of one or more of the other determinants.

Without loss of generality, we will assume that the difference in z to be studied relates to a difference over time. Denoting the value of z in the initial period 0 by z^0 and its value in the final period 1 as z^1, we can write

$$z^0 = x_1^0 x_2^0 ... x_n^0 \quad (2)$$

$$z^1 = x_1^1 x_2^1 ... x_n^1 \quad (3)$$

To decompose the change in z, two approaches can be chosen. First, the *ratios* between the left-hand sides and the right-hand sides of equations (3) and (2) provide the starting point for a multiplicative decomposition form. In an input-output context, this approach was rather recently introduced by Dietzenbacher *et al.* (2000). We will not pursue this approach in this paper, however. Instead, we will focus on the probably more popular additive decomposition form, which is based on the *differences* between the left-hand sides and the right-hand sides of equations (3) and (2). We obtain:

$$\Delta z = z^1 - z^0 = x_1^1 x_2^1 ... x_n^1 - x_1^0 x_2^0 ... x_n^0 \quad (4)$$

The objective of additive decomposition analyses is now to express the value of the left-hand side as the sum of the respective effects of every determinant x_i:

$$\Delta z = \Delta x_1 \text{ effect} + \Delta x_2 \text{ effect} + ... + \Delta x_n \text{ effect} \quad (5)$$

To explain the nature of the non-uniqueness problem that emerges, we rely on the case in which n=2. For notational convenience, we will denote the exogenous variables by x and y. Hence, we have:

$$z = xy \quad (6)$$

and

$$\Delta z = z^1 - z^0 = x^1 y^1 - x^0 y^0 \quad (7)$$

Now, we can obtain the equivalent of equation (5) by adding and subtracting $x^0 y^1$ in (7), obtaining:

$$\Delta z = x^1 y^1 - x^0 y^0 + x^0 y^1 - x^0 y^1 = \left(x^1 - x^0\right) y^1 + x^0 \left(y^1 - y^0\right) \tag{8}$$

and:

$$\Delta z = \Delta x y^1 + x^0 \Delta y \tag{9}$$

The first term on the right side of (9) represents the effect of changes in x to the actual change in z, and the second term quantifies the contribution of changes in variable y. The problem arises because different contributions could have been obtained if we had added and subtracted $x^1 y^0$ in (7) instead of $x^0 y^1$. In this case, we would have obtained:

$$\Delta z = \Delta x y^0 + x^1 \Delta y \tag{10}$$

The contributions of changes in x and y as obtained by expressions (9) and (10) can differ quite a bit and choosing one of them is an arbitrary decision.[4] As a pragmatic solution, authors have traditionally applied average solutions of expressions (9) and (10). As Dietzenbacher & Los (1998) pointed out, this is equal to using midpoint weights if and only if two determinants are discerned.

$$\Delta z = \Delta x y^{(½)} + x^{(½)} \Delta y \tag{11}$$

where,

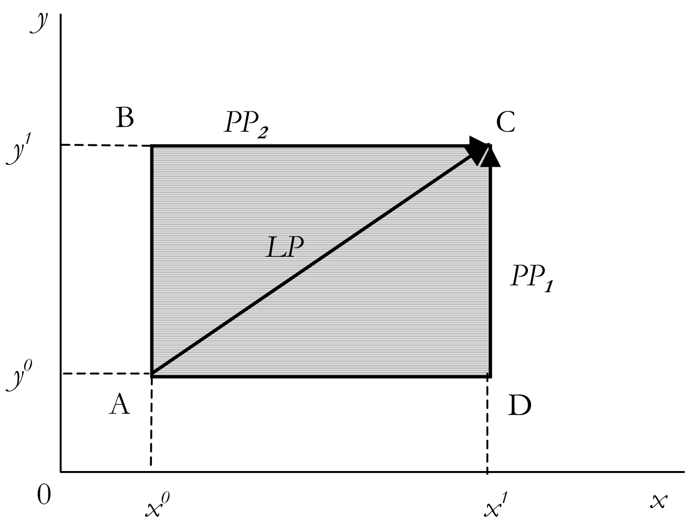
$$x^{(½)} = \frac{x^0 + x^1}{2} \quad \text{and} \quad y^{(½)} = \frac{y^0 + y^1}{2} \tag{12}$$

---

4 We only consider "exhaustive" decomposition forms, which implies that the full effect is attributed to changes in the exogenous determinants. An example of a "non-exhaustive" or "approximate" (Dietzenbacher & Los, 1998) decomposition form is $\Delta z = \Delta x y^0 + x^0 \Delta y + \Delta x \Delta y$. The last term is often labelled the "interaction effect". In some cases, approximate forms may be preferred over exhaustive forms, for example if a clear economic interpretation can be given to the interaction term. If $n>2$, however, approximate decompositions will contain a number of interaction terms, for which no straightforward interpretation is available. In such cases, we feel that exhaustive decomposition forms are most appropriate.

This discussion is graphically summarized by Figure 1, which was originally proposed by Sun (1998). In period 0, the value of z is represented by the small lower left rectangle (Oy^0Ax^0). In period 1, it is given by the surface of the larger rectangle Oy^1Cx^1. It is undisputable that the parts of the change given by the rectangles x^0ADx^1 and y^0y^1BA should be attributed to the growth of x and y, respectively. The whole issue is about the treatment of the upper right rectangle (ABCD), the interaction effect. Equation (9) suggests to attribute it completely to the change in x, whereas equation (10) would attribute it completely to the change in y. Consequently, the contributions for x and y obtained by both expressions can imply remarkable differences. The actual size of the difference depends on the size of the interaction term. The larger its size, the larger the variability among the results.

Figure 1. Polar and straight-line paths



The specification of a temporal path for the determinants implies a particular decomposition form to split-up the interaction term. We will get back to the issue of temporal paths in much more detail in the next section. For now, it should be noted that path PP_1 would mean that the effect of determinant x would be $\Delta x y^0$, and the effect of determinant y would be $x^1 \Delta y$. If we suppose that the temporal path between the initial and the final period is path PP_2, the respective contributions for determinants x and y would be $\Delta x y^1$ and $x^0 \Delta y$. Taking the average of these two alternative paths would imply an equal division of the interaction rectangle. It can easily be seen that taking the midpoint weights would yield an identical result. This result is also attained by Sun's (1998) method, which amounts to attribute halves of the interaction effect to the effects of changes in the two determinants. This amounts to drawing a straight line (*LP*) from (x^0, y^0) to (x^1, y^1).[5]

In the general case, in which z is the product of n determinants, the number of possible basic decompositions such as those corresponding to PP_1 and PP_2 is increased, now being equal to the number of possible permutations for n variables. Therefore, $n!$ forms could be obtained to decompose the change Δz. Specific cases among these are

$$\Delta z = \Delta x_1 x_2^0 ... x_n^0 + x_1^1 \Delta x_2 ... x_n^0 + ... + x_1^1 x_2^1 ... \Delta x_n \tag{13}$$

$$\Delta z = \Delta x_1 x_2^1 ... x_n^1 + x_1^0 \Delta x_2 ... x_n^1 + ... + x_1^0 x_2^0 ... \Delta x_n \tag{14}$$

These expressions are usually called "polar decompositions" (Dietzenbacher & Los, 1998), because the expressions for the effects are characterized by identical indexes for all determinants on both the left hand-side and right hand-side of the Δx_i factor.[6] The absence of uniqueness in the solutions leads to the arbitrary choice for one of the $n!$ possibilities, or alternatively one could obtain an average solution. As Dietzenbacher & Los (1998) showed, the average of the two polar decompositions is usually very close to the average taken over all $n!$ forms. They also show that a midpoint weighted formula is not exhaustive if $n>2$.

In the next sections, we will study the main features of a general method of decomposition that overcomes many of the limitations of the SDA approaches discussed so far. It allows us to obtain non-arbitrary solutions to measure the effects of the determinants of a change.

3. The Path Based Approach

In this section, a framework for an alternative decomposition method will be sketched. It builds on earlier work by Hoekstra & Van den Bergh (2002) and in particular Harrison *et al.* (2000), who introduced the basics of what we will call the Path Based (PB) approach. The alternative setup starts from the premise that both the value of z and the value of the determinants x_i have changed continuously over time, between time 0 and time 1. Hence, we can write:

$$z(t) = x_1(t) x_2(t) ... x_n(t) \tag{15}$$

and, assuming differentiability of each $x_i(t)$ an infinitesimal change in z can be expressed as

---

5 Sun's (1998) straight line can be considered as a special case of the continuous-time approach we will discuss below. Sun himself refers to his solution as the implication of a "jointly created and equally distributed" principle (Sun, 1998, p. 88).

6 In fact, this property is also fulfilled by the decomposition forms corresponding to PP_1 and PP_2. We will therefore denote such paths as "polar paths" (PP).

$$dz = \frac{\partial z}{\partial x_1}\frac{dx_1}{dt}dt + \ldots + \frac{\partial z}{\partial x_n}\frac{dx_n}{dt}dt \qquad (16)$$

Finally, the total change in z can be expressed as the sum of all the infinitesimal changes between time 0 and time 1:

$$\Delta z = \int_{t=0}^{t=1} \frac{dz}{dt}dt = \int_{t=0}^{t=1} \sum_{i=1}^{n} \frac{\partial z}{\partial x_i}\frac{dx_i}{dt}dt \qquad (17)$$

The effects of the determinants x_i can now be written as:

$$\Delta x_i \text{effect} = \int_{t=0}^{t=1} \frac{\partial z}{\partial x_i}\frac{dx_i}{dt}dt = \int_{t=0}^{t=1} \prod_{j\neq i}^{n} x_j \frac{dx_i}{dt}dt \qquad (18)$$

Equation (18) shows that the derivatives of the determinants x_i to time t play an important role in the size of the effects attributed to changes in these determinants. Consequently, the choice of the functional forms of the functions $x_i(t)=f_i(t)$, or in other words, the specification of the temporal paths that variables follow between initial and final periods, can have a big impact on the measurement of their effects that together add up to the variation in z.

Harrison *et al.* (2000) proposed the solution arrived at by assuming straight-line paths of the variables x_i:

$$x_i(t) = x_i^0 + \left(x_i^1 - x_i^0\right)t = x_i^0 + \Delta x_i t \qquad (19)$$

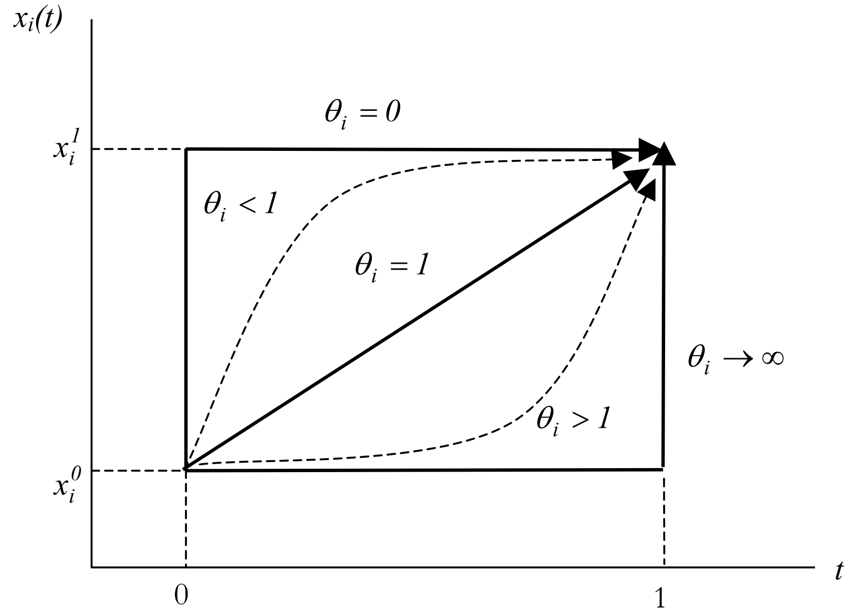
In the case of two determinants, this procedure amounts to attributing half of the interaction effect to the first determinant and the other half to the second determinant. Actually, this approach yields the same solution as Sun's (1998) 'equal shares' method. Empirical values of the variables x and y at t=0.5 might however be such that the straight line assumption is very unlikely to be tenable. If, for example, $x(0.5)$ would be very close to $x(0)$ and $y(0.5)$ would be very close to $y(1)$, it would clearly be preferable to opt for attribution of the interaction term to the Δx-effect (path PP_2 in Figure 1). In this paper, we propose a method to take such information explicitly into account in attributing parts of the interaction effects to the effects of the respective determinants.

The methodological innovation we propose is to relax the strict assumption of a straight line, by considering more flexible forms for the functions $f_i(t)$. In order to preserve possibilities to estimate the parameters that characterize the time-paths of the variables, we choose to consider a specific class of monotonic functions without inflexion points:

$$x_i(t) = x_i^0 + \Delta x_i t^{\theta_i};\ \forall\, \theta_i > 0 \tag{20}$$

Obviously, the temporal path of x_i will be a straight line if θ_i equals 1. If this holds for all i (i=1, ..., n), the solution obtained by the method introduced here will be identical to Harrison's et al. (2000) solution. Figure 2 indicates what the path for x_i looks like if θ_i takes on a values different from 1.

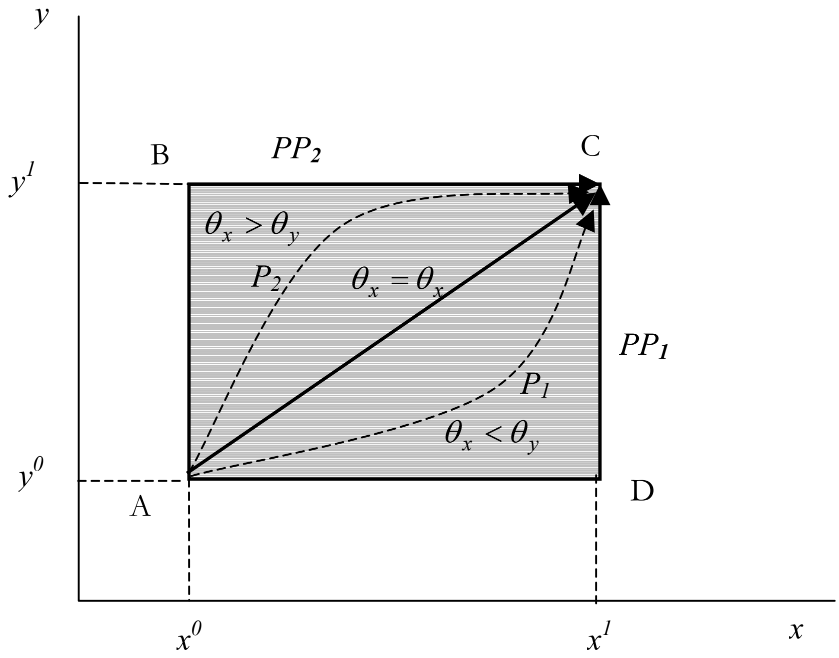
Figure 2. Several temporal paths for factor $\mathbf{x}_i$



As can be seen from Figure 2, the class of paths considered contains all possible monotonic paths for x_0 to x_1 that do not have inflexion points. This is a limitation for sure. An important category of paths not covered by our class of paths are those that contain values that are below the initial value or exceed the final value (assuming, without loss of generalization that x_1 is larger than x_0). However, by plotting a diagram for two determinants comparable to Figure 1, we can show that the class of time paths implied by the still relatively simple expression in equation (20) comprises a nicely defined set of time paths (see Figure 3). The "polar" paths PP_1 ($\theta_x/\theta_y \to 0$) and PP_2 ($\theta_x/\theta_y \to \infty$), and the straight-line path (θ_x/θ_y=1) are included as special cases of this general class.[7] P_1 and P_2 are intermediate cases.

[7] See Fernández (2004, pp. 29-32) for formal analysis proving these points for the case of two determinants.

Figure 3. Generalized monotonic temporal paths



The basic idea is that the specific path implied by the parameter values θ_i determines the shares of the interaction effect that is attributed to the distinct determinants. In a situation like the one represented by P_1 in Figure 3, a larger part of the interaction effect is attributed to determinant y than if a situation better reflected by P_2 would occur. In the next section, we will propose a methodology to estimate the values of θ_i, which allows us to decide whether a path like P_1 or P_2 (or PP_2, for that matter) describes the real-world trends best. Before we turn to that important issue, we first express the general idea outlined so far in more analytical terms.

For the most general case in which a change in z is decomposed into the effects of n determinants x_i (see equation (15)), the expression for the respective contributions for any possible set of n time paths was already given in equation (18). Substituting the more specific temporal paths assumed in equation (20) into equation (18), we can write

$$\Delta x_i \text{ Effect} = \int_{t=0}^{t=1} \prod_{j \neq i}^{n} x_j \frac{dx_i}{dt} dt = \left[\prod_{j<i}^{i-1} x_j^0 \right] \Delta x_i \left[\prod_{j>i}^{n} x_j^0 \right] + \tag{21a}$$

$$+ \sum_{j \neq i}^{n} \left[\frac{\theta_i}{\theta_i + \theta_j} \prod_{k<i}^{i-1} x_k^0 \Delta x_i \prod_{i<k<j}^{j-1} x_k^0 \Delta x_j \prod_{k>j}^{n} x_k^0 \right] + \tag{21b}$$

$$+ \sum_{j \neq i}^{n} \sum_{l \neq j,i}^{n} \left[\frac{\theta_i}{\theta_i + \theta_j + \theta_l} \prod_{k<i}^{i-1} x_k^0 \Delta x_i \prod_{i<k<j}^{j-1} x_k^0 \Delta x_j \prod_{j<k<l}^{l-1} x_k^0 \Delta x_l \prod_{k>l}^{n} x_k^0 \right] + \tag{21c}$$

$$\cdots$$

$$+\frac{\theta_i}{\sum_{j=1}^{n}\theta_j}\left[\prod_{j=1}^{n}\Delta x_j\right] \tag{21d}$$

The first term in this sum shows the smallest contribution for determinant x_i, which is given by its growth Δx_i weighted by the initial values of the other variables.[8] It does not contain any part of the interaction effects. The remaining terms show a set of interaction effects between the growth of groups of determinants, also weighted by the initial values of the remaining determinants. The distribution of these joint effects among effects of determinants clearly depends on the θ_i values. Multiple joint effects between the determinants exist. More specifically, there are $\binom{n-1}{1}$ possibilities of interaction between x_i and each one of the remaining $n-1$ determinants, $\binom{n-1}{2}$ terms measuring the joint effect of x_i with groups of $n-2$ determinants, etc. In general, in the expression for the effect of x_i there will be $\binom{n-1}{k}$ terms for the joint effects with groups of *k* determinants. The last terms (in equation 21d), shows the part of the joint contribution of all the determinants to the interaction effect attributed to x_i.

The importance of the values of the θ_i parameters for the measurement of the determinant's contributions is clear from equation (21). The higher the value of θ_i in comparison to the remaining θ_j, the greater the portions of the interaction effects attributed to x_i and, thus, the greater its contribution to the whole change in variable *z*. To illustrate this idea, it is helpful to give extreme values to a parameter θ_i. Let us suppose firstly that θ_i tends to its minimum value, *i.e.* we are supposing that it is very close to zero. In this case we obtain:

$$\lim_{\theta_i\to 0}\Delta x_i \text{ effect} = \left[\prod_{j<i}^{i-1}x_j^0\right]\Delta x_i\left[\prod_{j>i}^{n}x_j^0\right] = x_1^0 x_2^0 \ldots x_{i-1}^0 \Delta x_i x_{i+1}^0 \ldots x_n^0 \tag{22}$$

This would be the case when the effect of changes in variable x_i is at its smallest, because we are weighting it by the remaining determinants at their initial values. It should be noted that equation (22) is one of the $n!$ feasible solutions obtained by SDA. The opposite situation will happen if we suppose that parameter θ_i has a much higher value than the rest of parameters θ_j. Then, the contribution of x_i will be:

$$\lim_{\theta_i\to\infty}\Delta x_i \text{ effect} = x_1^1 x_2^1 \ldots x_{\mathrm{i-1}}^1 \Delta x_i x_{\mathrm{i+1}}^1 \ldots x_{\mathrm{n}}^1 \tag{23}$$

[8] For the sake of simplicity, let us hereafter suppose a situation in which $\Delta x_i \geq 0$; $i = 1,\ldots,n$.

In such a case, the contribution of x_i to changes in variable z is as large as it can be, since we are weighting its variation by the remaining determinants measured at their final values. Between these two extreme situations there exists a infinite range of possible contributions for each determinant, which depend on the value of parameters θ_i. All solutions obtained by SDA techniques are included in this range.

As we mentioned before, it is nowadays common practice in SDA analyses to present averages over decomposition forms. The average over all *n!* decomposition forms could be obtained by the PB approach as well. If we would not have any information on the evolution of the determinants over time other than the initial and the final observation, it would be most plausible to assume that the temporal path parameters are equal to each other ($\theta_1 = \theta_2 = ... = \theta_n$). According to equation (21) we would find

$$\Delta x_i \text{ effect} = \int_{t=0}^{t=1} \prod_{j \neq i}^{n} x_j \frac{dx_i}{dt} dt = \left[\prod_{j<i}^{i-1} x_j^0 \right] \Delta x_i \left[\prod_{j>i}^{n} x_j^0 \right] + \tag{24a}$$

$$+ \sum_{j \neq i}^{n} \left[\frac{1}{2} \prod_{k<i}^{i-1} x_k^0 \Delta x_i \prod_{i<k<j}^{j-1} x_k^0 \Delta x_j \prod_{k>j}^{n} x_k^0 \right] + \tag{24b}$$

$$+ \sum_{j \neq i}^{n} \sum_{l \neq j,i}^{n} \left[\frac{1}{3} \prod_{k<i}^{i-1} x_k^0 \Delta x_i \prod_{i<k<j}^{j-1} x_k^0 \Delta x_j \prod_{j<k<l}^{l-1} x_k^0 \Delta x_l \prod_{k>l}^{n} x_k^0 \right] + \tag{24c}$$

$$\cdots$$

$$+ \frac{1}{n} \left[\prod_{j=1}^{n} \Delta x_j \right] \tag{24d}$$

The interaction effects are thus shared proportionally to the changes in the values of the determinants. This is identical to the solution proposed by Sun (1998) discussed in the previous section. In spite of the similarity between the numerical outcomes for the mean of the two polar decompositions only and the mean of all *n!* decompositions (Dietzenbacher & Los, 1998), the mean of the polar decompositions cannot be obtained by means of specifying values for θ_i in the above-mentioned PB approach.[9]

In the next sections we will turn to methods to infer on plausible values for θ_i, which allow us to apply equation (21) to interesting empirical problems.

[9] See Fernández (2004, pp. 36-39) for a proof.

4. Maximum Entropy Econometrics

In the previous section, we found that taking the mean contributions of all decomposition forms is the most reasonable solution to the non-uniqueness problem if the researcher has no information at all about the time paths of the determinants. In many cases, however, more information than the values of the determinants at t=0 and t=1 is available, for example about values of one or more of the determinants at intermediate points in time. Estimation of the parameters θ_i is generally not possible by means of classical econometric estimation procedures like least squares estimation. The amount of data is quite limited, which precludes the use of least squares estimation procedures based on limit theorems. Such procedures require at least more observations than parameters to be estimated, which is problematic in the input-output context studied here.

In this section, we will give an introduction to maximum entropy (ME) econometrics, a collection of tools that can be very convenient to use scarce additional information in producing estimates for the temporal path parameters θ_i.[10] To start with, let us assume that an event can have K possible outcomes E_1, E_2, ..., E_K with the respective distribution of probabilities $\mathbf{p} = p_1, p_2, ..., p_K$ such that $\sum_{i=1}^{K} p_i = 1$. Following the formulation of Shannon (1948), the entropy of this distribution $\mathbf{p}$ will be

$$H(\mathbf{p}) = -\sum_{i=1}^{K} p_i \ln p_i \tag{25}$$

which reaches its maximum when $\mathbf{p}$ is a uniform distribution ($p_i = \frac{1}{K}, \forall \mathrm{i} = 1,...,\mathrm{K}$). The entropy measure H indicates the 'uncertainty' of the outcomes of the event. If some information (*i.e.*, observations) is available, it can be used to estimate an unknown distribution of probabilities for a random variable x which can get values $\{x_1,...,x_K\}$.

Suppose that there are T observations $\{y_1, y_2,..., y_T\}$ available such that

$$\sum_{i=1}^{K} p_i f_t(x_i) = y_t, \ 1 \leq \mathrm{t} \leq \mathrm{T} \tag{26}$$

with $\{f_1(x), f_2(x),..., f_T(x)\}$ a set of known functions representing the relationships between the random variable x and the observed data $\{y_1, y_2,..., y_T\}$. In such a case, the ME principle can be applied to recover the unknown probabilities. This principle is based on the selection of the

---

10 See Kapur & Kesavan (1992) or Golan *et al.* (1996) for a detailed analysis of properties of the estimators obtained by means of these techniques.

probability distribution that maximizes equation (25) among all the possible probability distributions that fulfill (26). The following constrained maximization problem is posed:

$$\underset{\mathbf{p}}{Max}\ H(\mathbf{p}) = -\sum_{i=1}^{K} p_i \ln p_i \tag{27}$$

subject to:

$$\sum_{i=1}^{K} p_i f_t(x_i) = y_t,\ \forall t = 1,...,T$$

$$\sum_{i=1}^{K} p_i = 1$$

In this problem, the last restriction is just a normalization constraint that guarantees that the estimated probabilities sum to one, while the first T restrictions guarantee that the recovered distribution of probabilities is compatible with the data for all T observations. The Lagrangian function for problem (27) is

$$L = -\sum_{i=1}^{K} p_i \ln p_i + \sum_{t=1}^{T} \lambda_t \left[y_t - \sum_{i=1}^{K} p_i f_t(x_i) \right] + \mu \left[1 - \sum_{i=1}^{K} p_i \right] \tag{28}$$

and the corresponding estimates for the probabilities p_i are

$$\hat{p}_i = \frac{\exp\left[-\sum_{t=1}^{T} \hat{\lambda}_t f_t(x_i) \right]}{\sum_{i=1}^{K} \exp\left[-\sum_{t=1}^{T} \hat{\lambda}_t f_t(x_i) \right]},\ \forall i = 1,...,K \tag{29}$$

with λ_t the Lagrangian multipliers associated to the first T restrictions in the constrained maximization problem (27). It is important to note that even for T=1 (a situation with only one observation), the ME approach yields an estimate of the probabilities. Hence, in situations in which the number of observations is not large enough to apply econometrics based on limit theorems, this approach can be used to obtain robust estimates of unknown parameters.[11] A disadvantage of ME estimators is that comparisons of means and variances of estimators are not possible. Such comparisons are common practice in classical least squares and maximum likelihood econometrics.

---

[11] Golan *et al.* (1996, p. 12) contains a simple, classic example of this technique, the so called "dice problem".

For our current purposes, it is important that the above-sketched procedure can be generalized and extended to the estimation of unknown parameters for traditional linear models. Let us suppose that the problem at hand is the estimation of a linear model where a variable y depends on n explanatory variables x_i:

$$\mathbf{y} = \mathbf{X\theta} + \mathbf{e} \tag{30}$$

where $\mathbf{y}$ is a $(T \times 1)$ vector of observations for y, $\mathbf{X}$ is a $(T \times n)$ matrix of observations for the x_i variables, $\boldsymbol{\theta}$ is the $(n \times 1)$ vector of unknown parameters $\boldsymbol{\theta}' = (\theta_1,...,\theta_n)$ to be estimated, and $\mathbf{e}$ is a $(T \times 1)$ vector reflecting the random term of the linear model. For each θ_i, it will be assumed that there is some information about its $M \geq 2$ possible realizations by means of a 'support' vector $\mathbf{b}' = (b_1,...,b^*,...,b_M)$, the elements of which are symmetrically distanced around a central value $\theta_i = b^*$ (the prior expected value of the parameter), with corresponding probabilities $\mathbf{p}_i' = (p_{i1},..., p_{iM})$. 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 $\boldsymbol{\theta}$ can be written as

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdots \\ \theta_n \end{bmatrix} = \mathbf{Bp} = \begin{bmatrix} \mathbf{b}' & \mathbf{0} & . & \mathbf{0} \\ \mathbf{0} & \mathbf{b}' & . & \mathbf{0} \\ . & . & . & . \\ \mathbf{0} & \mathbf{0} & . & \mathbf{b}' \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \cdots \\ \mathbf{p}_n \end{bmatrix} \tag{31}$$

with $\mathbf{B}$ and $\mathbf{p}$ of dimensions (nxnM) and (nMx1), respectively. The value for each parameter is then given by

$$\theta_i = \mathbf{b}'\mathbf{p}_i = \sum_{m=1}^{M} b_m p_{im} \ ; \ \forall i = 1,...,n \tag{32}$$

For the random terms, a similar approach is chosen. To express the lack of information about the actual values contained in $\mathbf{e}$, we assume a distribution for each e_t, with a set of $J \geq 2$ values $\mathbf{v}' = (v_1,...,v_J)$ with respective probabilities $\mathbf{w}_t' = (w_{t1}, w_{t2},..., w_{tJ})$.[12] Hence, we can write

[12] Usually, the distribution for the errors is assumed symmetric and centered about 0, therefore $v_1 = -v_J$.

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_T \end{bmatrix} = \mathbf{Vw} = \begin{bmatrix} \mathbf{v'} & \mathbf{0} & . & \mathbf{0} \\ \mathbf{0} & \mathbf{v'} & . & \mathbf{0} \\ . & . & . & . \\ \mathbf{0} & \mathbf{0} & . & \mathbf{v'} \end{bmatrix} \begin{bmatrix} \mathbf{w_1} \\ \mathbf{w_2} \\ \dots \\ \mathbf{w_T} \end{bmatrix} \tag{33}$$

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

$$e_t = \mathbf{v'w_t} = \sum_{j=1}^{J} v_j w_{tj} \; ; \; \forall t = 1,...,T \tag{34}$$

And, consequently, equation (30) can be transformed into

$$\mathbf{y} = \mathbf{XBp} + \mathbf{Vw} \tag{35}$$

Now, the estimation problem for the unknown vector of parameters $\boldsymbol{\theta}$ is reduced to the estimation of $n+T$ probability distributions, and the following maximization problem (similar to problem (27)) can be solved to obtain these estimates:

$$\underset{\mathbf{p},\mathbf{w}}{Max}\, H(\mathbf{p},\mathbf{w}) = -\sum_{i=1}^{n}\sum_{m=1}^{M} p_{im} \ln(p_{im}) - \sum_{t=1}^{T}\sum_{j=1}^{J} w_{tj} \ln(w_{tj}) \tag{36}$$

subject to:

$$\sum_{i=1}^{n}\sum_{m=1}^{M} x_{it} b_m p_{im} + \sum_{j=1}^{J} v_j w_{tj} = y_t, \; \forall \mathrm{t} = 1,...,T$$

$$\sum_{m=1}^{M} p_{im} = 1, \; \forall i = 1,...,n$$

$$\sum_{j=1}^{J} w_{tj} = 1, \; \forall t = 1,...,T$$

By solving the associated Lagrangian function, we find

$$\hat{p}_{im} = \frac{\exp\left[-\sum_{t=1}^{T} \hat{\lambda}_t x_{it} b_m\right]}{\sum_{m=1}^{M} \exp\left[-\sum_{t=1}^{T} \hat{\lambda}_t x_{it} b_m\right]}, \; \forall i = 1,...,n; \; \forall m = 1,...,M \tag{37}$$

$$\hat{w}_{tj} = \frac{\exp\left[-\sum_{t=1}^{T}\hat{\lambda}_t v_j\right]}{\sum_{j=1}^{J}\exp\left[-\sum_{t=1}^{T}\hat{\lambda}_t v_j\right]}, \ \forall t = 1,\dots,T; \ \forall j = 1,\dots,J \tag{38}$$

Finally, these estimated probabilities allow us to obtain estimations for the unknown parameters.[13] The estimated value of θ_i will be:[14,15]

$$\hat{\theta}_i = \sum_{m=1}^{M}\hat{p}_{im}b_m , \ \forall i = 1,\dots,n \tag{39}$$

This approach can be applied to the decomposition problem studied in the previous section, since limited additional information would enable us to obtain estimates of the parameters that determine the contribution of each determinant to the total change that has actually been observed. In other words, non-arbitrary solutions to the decomposition problem could be obtained. In the next section several situations with availability of various types of additional data will be considered, as well as the way to estimate the effects of the factors to the total change Δz using this technique.

5. Incorporating Additional Information in SDA

In this section we will first suppose a scenario in which we have some additional observations for intermediate periods. A "dynamic SDA" in the more traditional sense is not possible, however, since we suppose that these intermediate observations are only available for a few of the *n*

13 Golan *et al.* (1996, Chapter 6) show that these estimators are consistent and asymptotically normal. In Golan *et al.* (1996, Chapter 7) the finite sample behavior of the ME estimators is numerically compared to traditional least squares and maximum likelihood estimators. In experimental samples with limited data, the ME estimators are found to be superior.

14 The construction of the vector **b** is based on the researcher's prior knowledge (or beliefs) about the parameter. Sometimes, the choice of minimum and maximum values b_1 and b_M is quite obvious, but in other cases a 'natural' choice does not exist. In such a situation, it will not be possible to obtain an accurate solution to the estimation problem if the actual parameter value is out of the fixed range, say $\theta_i > b_M$. Therefore, one should be careful in choosing the maximum and minimum values of **b**. Golan *et al.* (1996, chapter 8) devote more attention to consequences of choices concerning the elements of the vector **b**. An almost universal result is that wider bounds can be used without substantial consequences for the characteristics of the estimators.

15 Fernández (2004, pp. 69) proves that the solution of the constrained maximization problem (36) without additional information yields estimates equal to the expected value b^* of the prior distribution.

factors.[16] To assess the contribution of factor x_i, equation (20) will be used again, but in a slightly different form. It contains a stochastic component ε_{it} that allows x_i to diverge from the deterministic path that we would like to estimate[17]

$$x_i(t) = x_i^0 + \Delta x_i t^{\theta_i} e^{\varepsilon_{it}} \tag{40}$$

Defining $g_i(t) = x_i(t) - x_i^0$ and taking logarithms, we have:

$$\ln\left(\frac{g_i(t)}{\Delta x_i}\right) = \theta_i \ln(t) + \varepsilon_{it}\text{, or}$$
$$x_i^*(t) = \theta_i t^* + \varepsilon_{it} \tag{41}$$

Equation (41) is a linear model with one parameter to be estimated. Hence, it is possible to apply the Maximum Entropy estimation technique for linear relationships analyzed in the previous section, and (41) can be written as

$$x_i^*(t) = \sum_{m=1}^{M} b_{i\,m} p_{i\,m} t^* + \sum_{j=1}^{J} v_j w_{t\,j} \tag{42}$$

which ends up as a constraint in the following maximization problem

$$\underset{\mathbf{p},\mathbf{w}}{Max}\, H(\mathbf{p},\mathbf{w}) = -\sum_{m=1}^{M} p_{i\,m} \ln(p_{i\,m}) - \sum_{t=1}^{T}\sum_{j=1}^{J} w_{t\,j} \ln(w_{t\,j}) \tag{43}$$

subject to:

$$x_i^*(t) = \sum_{m=1}^{M} b_{i\,m} p_{i\,m} t^* + \sum_{j=1}^{J} v_j w_{t\,j},\ \ \forall t = 1,\ldots,T$$

$$\sum_{m=1}^{M} p_{im} = 1$$

[16] The term "dynamic SDA" was inspired by the related term "dynamic shift-share analysis" proposed by Barff & Knight III (1988). If observations for x_1, x_2, ..., x_n would be available for a period s ($0<s<1$), dynamic SDA would amount to decomposing z_s-z_0 and z_1-z_s in the classic way outlined in Section 2, and subsequently adding results for the corresponding effects in the two decompositions to obtain the contributions for z_1-z_0. A discussion of transitivity problems in this approach is beyond the scope of this paper.

[17] We assume that $\varepsilon_{it} = 0$ in the final period. This ensures that $x_i(t)$ has value x_i^1 in this period.

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

According to equation (39), solving this problem yields an estimate for parameter θ_i. For factors x_i for which there is no additional information, the estimates $\hat{\theta}_j$ should equal 1 to resemble the linear path. Hence, the central value b^* should be set to 1. Upon having obtained estimates for θ_1, θ_2, ..., θ_n, substitution of these values and the observations for $x_1, x_2, \dots, x_n$ in equation (21) yields the estimated respective contributions of changes in the determinants.

Another situation in which additional information can be incorporated in SDA emerges if we do not have intermediate observations for the factors involved in the decomposition problem itself, but have such observations for other variables highly correlated with them. In the next section we will discuss such an empirical case.

Assume that there is another variable $r_i(t)$, which is correlated to a determinant $x_i(t)$ and directly observable for at least one intermediate point between the initial and final periods (as opposed to $x_i(t)$ itself). In such a situation, using the initial and final values for both variables, a linear model such as

$$x_{ik} = \beta_0 + \beta_1 r_{ik} + \varepsilon_k \quad (k=1, \dots, K, K+1, \dots, 2K) \tag{44}$$

can be estimated, where ε_k is a random term with the usual characteristics and K is the dimension of determinant $x_i(t)$.[18] This model can be employed to estimate the $x_i(t)$ values from $r_i(t)$ values. If there are T observations for variable $r_i(t)$, T estimates for $x_i(t)$ can be obtained as:

$$\hat{x}_i(t) = \hat{\beta}_0 + \hat{\beta}_1 r_i(t); \; t = 1,\dots,T \tag{45}$$

and it would be possible to include these $\hat{x}_i(t); \; t = 1,\dots,T$ values as additional information in a ME program to estimate the parameter θ_i. It should be kept in mind that the accuracy of these estimates is positively related to the strength of the correlation. To take this into account, we propose to construct confidence intervals:

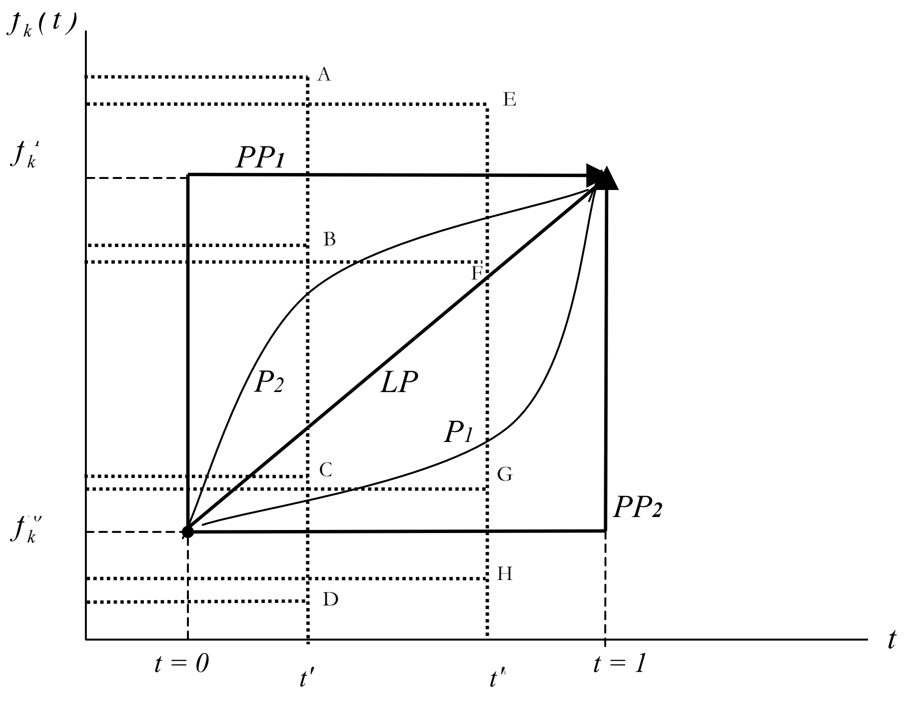
$$\left[\hat{x}_i(t) \pm a S_{\hat{x}}\right] \tag{46}$$

[18] If $x_i(t)$ and $r_i(t)$ are vectors of dimension Kx1, for example, equation (44) can be estimated using K observations for the initial period and K observations for the final period, adding up to a total of $2K$ observations. This implies that this approach is only feasible if the determinants have at least dimension 2x1.

where $S_{\hat{x}}$ is the standard error of the estimated value $\hat{x}_i(t)$ and *a* is the *t*-value that corresponds to the confidence level 1-α. The lower and upper values of such intervals, say $\hat{x}_i^{low}(t)$ and $\hat{x}_i^{high}(t)$, will be used to estimate intervals for the parameters θ_i. The lower and upper bounds ($\hat{\theta}_i^{low}$ and $\hat{\theta}_i^{high}$, respectively) are found by including $\hat{x}_i^{low}(t)$ and $\hat{x}_i^{high}(t)$ as additional information in ME programs. More specifically, these values show up as the left hand side variables in the first constraint in maximization problem (43). Since it is possible to estimate intervals for the parameters, it will be possible to compute intervals for the contributions of the determinants, too. Note that the higher correlation between determinant $x_i(t)$ and variable $r_i(t)$, the greater the expected accuracy of the estimate of the parameter, which implies narrower confidence intervals for the contributions of changes in the determinants.

The use of both kinds of additional information in the framework outlined above can cause nontrivial problems if the information is rather unlikely to be generated by a time path belonging to the class of paths defined by equation (40). This happens if observations for intermediate periods rule out a monotonic path. In Figure 3, such points are located above or below the rectangle. We deal with such observations by fitting the most appropriate monotonic paths. Figure 4 depicts all possibilities if two intermediate observations are available.

Figure 4. Estimated temporal paths with intermediate observations



For intermediate period *t'* observations for this determinant can be categorized as A, B, C or D, depending on whether they are above or below the linear path and inside or outside the rectangle. In the same vein, we have E, F, G or H for intermediate period *t''*. If the two observations are both like B, C, F and G, no problems are encountered. If A and E are observed, the closest monotonic path is PP_1, which corresponds to θ_i=0. If D and H are observed, PP_2 is most appropriate and $\theta_i=\infty$.[19] If A (above the rectangle) and H (below the rectangle) are observed, we opt for the linear path, since it is the average of PP_1 (implied by A) and PP_2 (implied by H). If points like B and E or B and H are observed, there will be an observation inside the rectangle (B) and another one in the outside (E or H). In such cases, to obtain valid estimates of the parameter it will be assumed that points E or H are not outside the rectangle but just on the border of the rectangle (given by PP_1 and PP_2 respectively). The same procedure is applied in situations with observations like A and F or D and F.

It should be noted that the two situations depicted concerning availability of additional information do not offer an exhaustive enumeration of all possibilities for incorporating such information into decomposition problems. The important issue is that the flexibility of this estimation method allows including information even if there were not direct observations of the factors appearing in the decomposition problem. If there is some kind of knowledge about the behavior of other variables that are somehow related to these factors, this information can be used to obtain estimates of the parameters.

6. Illustration: Sectoral Labor Cost Growth in the Spanish Economy

We apply the techniques developed in the previous sections to study the contributions of three determinants to changes in real sectoral labor costs in Spain, over the period 1980-1994. It should be emphasized that the aim of this section is not so much to provide a "deep" analysis of the dynamics of Spanish labor costs, but rather to provide an illustration of the methods proposed in this paper. The required data were taken from 21-sector input-output tables for these years, expressed in prices of 1986. The intermediate blocks of the tables contain domestic deliveries only. Appendix A contains detailed information about how we treated the basic data to arrive at the data used in the analysis outlined below.

Our starting point is an input-output model that expresses the vector of sectoral labor costs **c** as the product of three factors, i.e. labor costs per unit of gross output **u** (included as a diagonal matrix), the Leontief inverse matrix **L** and the vector of final demands **f**.

$$\mathbf{c} = \hat{\mathbf{u}}\mathbf{L}\mathbf{f} \tag{47}$$

and the objective is to decompose the total change $\Delta\mathbf{c}$ into the following three components:

---

[19] If $\theta_i=\infty$, a "very big" value must be inserted in equation (21) to obtain numerical results. In the empirical application described in the next section, we used the value 10^{20} in such cases.

$$\mathbf{c}^{94} - \mathbf{c}^{80} = \boldsymbol{\Delta}\mathbf{c} = \boldsymbol{\Delta}\mathbf{u} \text{ effect} + \boldsymbol{\Delta}\mathbf{L} \text{ effect} + \boldsymbol{\Delta}\mathbf{f} \text{ effect} \tag{48}$$

We assume the following temporal paths for the elements of the factors:

$$u_k(t) = u_k^{80} + \Delta u_k t^{\theta_{u_k}} \text{ ; } k=1, ..., 21 \tag{49a}$$

$$l_{kj}(t) = l_{kj}^{80} + \Delta l_{kj} t^{\theta_{l_{kj}}} \text{ ; } k=1, ..., 21; j=1, ..., 21 \tag{49b}$$

$$f_k(t) = f_k^{80} + \Delta f_k t^{\theta_{f_k}} \text{ ; } k=1, ..., 21 \tag{49c}$$

According to equation (21), the contributions of changes in elements of **u**, **L** and **f** to the changes in **c** can be written as[20]

$$\Delta\mathbf{u} \text{ effect} = \Delta\hat{\mathbf{u}}\mathbf{L}^{\mathbf{80}}\mathbf{f}^{\mathbf{80}} + \left[\boldsymbol{\Theta}_{u+L}^{u} \circ \Delta\hat{\mathbf{u}}\Delta\mathbf{L}\right]\mathbf{f}^{80} + \left[\boldsymbol{\Theta}_{u+f}^{u} \circ \Delta\hat{\mathbf{u}}\mathbf{L}^{80}\right]\Delta\mathbf{f} + \left[\boldsymbol{\Theta}_{u+L+f}^{u} \circ \Delta\hat{\mathbf{u}}\Delta\mathbf{L}\right]\Delta\mathbf{f} \tag{50a}$$

$$\Delta\mathbf{L} \text{ effect} = \hat{\mathbf{u}}^{80}\Delta\mathbf{L}\mathbf{f}^{\mathbf{80}} + \left[\boldsymbol{\Theta}_{u+L}^{L} \circ \Delta\hat{\mathbf{u}}\Delta\mathbf{L}\right]\mathbf{f}^{80} + \left[\boldsymbol{\Theta}_{L+f}^{L} \circ \hat{\mathbf{u}}^{80}\Delta\mathbf{L}\right]\Delta\mathbf{f} + \left[\boldsymbol{\Theta}_{u+L+f}^{L} \circ \Delta\hat{\mathbf{u}}\Delta\mathbf{L}\right]\Delta\mathbf{f} \tag{50b}$$

$$\Delta\mathbf{f} \text{ effect} = \hat{\mathbf{u}}^{80}\mathbf{L}^{80}\Delta\mathbf{f} + \left[\boldsymbol{\Theta}_{u+f}^{f} \circ \Delta\hat{\mathbf{u}}\mathbf{L}^{80}\right]\Delta\mathbf{f} + \left[\boldsymbol{\Theta}_{L+f}^{f} \circ \hat{\mathbf{u}}^{80}\Delta\mathbf{L}\right]\Delta\mathbf{f} + \left[\boldsymbol{\Theta}_{u+L+f}^{f} \circ \Delta\hat{\mathbf{u}}\Delta\mathbf{L}\right]\Delta\mathbf{f} \tag{50c}$$

with the matrices $\boldsymbol{\Theta}$ defined as

$$\boldsymbol{\Theta}_{u+L}^{u} \equiv \begin{bmatrix} \dfrac{\theta_{u_1}}{\theta_{u_1} + \theta_{l_{11}}} & \cdots & \dfrac{\theta_{u_1}}{\theta_{u_1} + \theta_{l_{1K}}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{u_K}}{\theta_{u_K} + \theta_{l_{K1}}} & \cdots & \dfrac{\theta_{u_K}}{\theta_{u_K} + \theta_{l_{KK}}} \end{bmatrix} \tag{51a}$$

$$\boldsymbol{\Theta}_{u+f}^{u} \equiv \begin{bmatrix} \dfrac{\theta_{u_1}}{\theta_{u_1} + \theta_{f_1}} & \cdots & \dfrac{\theta_{u_1}}{\theta_{u_1} + \theta_{f_K}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{u_K}}{\theta_{u_K} + \theta_{f_1}} & \cdots & \dfrac{\theta_{u_K}}{\theta_{u_K} + \theta_{f_K}} \end{bmatrix} \tag{51b}$$

[20] The symbol $\circ$ indicates element-by-element (Hadamard) multiplication.

$$\boldsymbol{\Theta}^{u}_{u+L+f} \equiv \begin{bmatrix} \dfrac{\theta_{u_1}}{\theta_{u_1}+\theta_{l_{11}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{u_1}}{\theta_{u_1}+\theta_{l_{1K}}+\theta_{f_K}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{u_K}}{\theta_{u_K}+\theta_{l_{K1}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{u_K}}{\theta_{u_K}+\theta_{l_{KK}}+\theta_{f_K}} \end{bmatrix} \tag{51c}$$

$$\boldsymbol{\Theta}^{L}_{u+L} \equiv \begin{bmatrix} \dfrac{\theta_{l_{11}}}{\theta_{u_1}+\theta_{l_{11}}} & \cdots & \dfrac{\theta_{l_{1K}}}{\theta_{u_1}+\theta_{l_{1K}}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{l_{K1}}}{\theta_{u_1}+\theta_{l_{K1}}} & \cdots & \dfrac{\theta_{l_{KK}}}{\theta_{u_K}+\theta_{l_{KK}}} \end{bmatrix} \tag{51d}$$

$$\boldsymbol{\Theta}^{L}_{L+f} \equiv \begin{bmatrix} \dfrac{\theta_{l_{11}}}{\theta_{l_{11}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{l_{1K}}}{\theta_{l_{1K}}+\theta_{f_K}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{l_{K1}}}{\theta_{l_{K1}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{l_{KK}}}{\theta_{l_{KK}}+\theta_{f_K}} \end{bmatrix} \tag{51e}$$

$$\boldsymbol{\Theta}^{L}_{u+L+f} \equiv \begin{bmatrix} \dfrac{\theta_{l_{11}}}{\theta_{u_1}+\theta_{l_{11}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{l_{1K}}}{\theta_{u_1}+\theta_{l_{1K}}+\theta_{f_K}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{l_{K1}}}{\theta_{u_K}+\theta_{l_{K1}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{l_{KK}}}{\theta_{u_K}+\theta_{l_{KK}}+\theta_{f_K}} \end{bmatrix} \tag{51f}$$

$$\boldsymbol{\Theta}^{f}_{u+f} \equiv \begin{bmatrix} \dfrac{\theta_{f_1}}{\theta_{u_1}+\theta_{f_1}} & \cdots & \dfrac{\theta_{f_K}}{\theta_{u_1}+\theta_{f_K}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{f_1}}{\theta_{u_K}+\theta_{f_1}} & \cdots & \dfrac{\theta_{f_K}}{\theta_{u_K}+\theta_{f_K}} \end{bmatrix} \tag{51g}$$

$$\boldsymbol{\Theta}^{f}_{L+f} \equiv \begin{bmatrix} \dfrac{\theta_{f_1}}{\theta_{l_{11}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{f_K}}{\theta_{l_{1K}}+\theta_{f_K}} \\ \vdots & \ddots & \vdots \\ \dfrac{\theta_{f_1}}{\theta_{l_{K1}}+\theta_{f_1}} & \cdots & \dfrac{\theta_{f_K}}{\theta_{l_{KK}}+\theta_{f_K}} \end{bmatrix} \tag{51h}$$

$$\mathbf{\Theta}^{f}_{u+L+f} \equiv \begin{bmatrix} \frac{\theta_{f_1}}{\theta_{u_1}+\theta_{l_{11}}+\theta_{f_1}} & \cdots & \frac{\theta_{f_K}}{\theta_{u_1}+\theta_{l_{1K}}+\theta_{f_K}} \\ \vdots & \ddots & \vdots \\ \frac{\theta_{f_1}}{\theta_{u_K}+\theta_{l_{K1}}+\theta_{f_1}} & \cdots & \frac{\theta_{f_K}}{\theta_{u_K}+\theta_{l_{KK}}+\theta_{f_K}} \end{bmatrix} \tag{51i}$$

As argued before, assuming that parameters θ are the same for all the factors (i.e., $\theta_{u_k} = \theta_{l_{kj}} = \theta_{f_k}$) for all *j* and *k* would yield Sun's (1998) solution. This would be a natural thing to do if no information would be available for the years in-between 1980 and 1994. To illustrate the techniques outlined in the previous sections, we will estimate some of the θ parameters by employing additional information of two sorts. First, we incorporate information about sectoral final demands for 1986 and/or 1990. Afterwards, we will employ information about wage payments per unit of output, which we consider as a variable correlated to the first determinant in the SDA, labor costs per unit of output.

In both cases, we had to decide on the values to be assigned to the a priori distributions contained in the support vector **b** (see equation 32) and the possible realizations for the random term in vectors **v** (see equation 33). The following vectors were used for all parameters throughout the empirical analyses below:[21]

b=[-5.0, -3.0, -1.0, 1.0, 3.0, 5.0, 7.0]' and **v**=[-0.01, -0.005, 0, 0.005, 0.01]'

6.a Inclusion of values of a determinant for an intermediate period

We will report two sets of results obtained by incorporating information on the volumes of sectoral final demands. First, we used observations for sectoral final demand in 1990 only. These levels were reported in the Spanish National Accounts (INE, 1999). We subsequently expressed them in 1986 prices by means of the deflation procedure discussed in Appendix A. Next, we applied the very same procedure, but incorporated sectoral final demand levels for 1986 (INE, 1987) and 1990, simultaneously. Consequently, Table 1 reports two sets of estimated θ_{f_i}, and two sets of associated decomposition results.

[21] Fernández (2004, p. 142-143) tested the assertion by Golan *et al.* (1996, p. 138) that the estimation results are generally not very sensitive to the choice of a particular set in the specific context of a path-based shift-share analysis. His results strongly confirmed this assertion by Golan *et al.*

Table 1: Decomposition results for case with additional information for one of the determinants.*

Sector	(1) Δ**c**[a]	(2) Δ**u** eff	(3) Δ**L** eff	(4) Δ**f** eff	(5) $\hat{\theta}_{f_i}$	(6) Δ**u** eff	(7) Δ**L** eff	(8) Δ**f** eff	(9) $\hat{\theta}_{f_i}$	(10) Δ**u** eff	(11) Δ**L** eff	(12) Δ**f** eff
1	24.2	-193.0	105.9	111.4	∞	-0.9	4.3	-5.7	∞	-3.3	-5.5	-0.2
2	-44.7	-132.6	35.2	52.7	0.00	-3.6	17.3	-20.5	0.00	-2.8	-10.9	2.2
3	53.2	-45.8	-76.0	175.0	∞	-5.4	-4.1	-3.2	∞	-6.7	-8.6	-5.4
4	42.5	-167.7	56.8	153.5	6.34	3.7	5.2	2.1	8.40	-9.3	-8.8	-7.0
5	261.5	-163.5	104.1	320.9	1.55	-4.3	4.0	-3.5	1.93	-5.2	-6.2	-0.5
6	130.0	-80.1	41.5	168.6	0.30	6.0	1.4	2.5	6.71	-9.7	-5.3	-3.5
7	26.5	-122.5	49.8	99.2	0.00	13.2	0.9	15.9	0.00	-1.8	-6.9	0.8
8	216.6	-267.8	72.1	412.4	0.79	5.6	5.3	2.7	1.25	-9.5	-7.8	-4.9
9	232.0	-403.8	32.9	602.9	0.92	2.2	8.5	1.0	1.49	-10.8	-8.8	-6.8
10	228.4	-15.5	98.6	145.3	0.55	3.1	3.6	-2.1	1.74	-5.4	-5.2	3.1
11	249.5	-141.3	58.4	332.4	0.43	9.7	6.2	3.1	0.67	-4.9	-9.0	-0.6
12	108.2	-74.4	86.5	96.1	1.90	-0.7	4.4	-4.5	1.78	-2.9	-6.6	4.1
13	-77.9	-156.7	39.4	39.4	0.00	-3.7	3.3	-18.2	0.00	-3.7	-5.3	-10.4
14	570.1	9.7	53.4	506.9	0.00	16.1	7.4	-1.1	0.90	-13.5	-8.5	1.3
15	527.0	-169.0	235.9	460.2	0.31	6.6	3.7	0.5	1.19	-7.7	-6.3	0.3
16	163.6	-36.7	35.7	164.6	0.23	8.7	3.5	1.2	0.78	-6.7	-5.9	-0.3
17	359.7	-463.5	506.6	316.6	1.71	-2.9	3.7	-10.1	1.42	-0.5	-6.5	11.2
18	105.8	83.5	-111.8	134.1	1.89	0.5	-0.5	-0.8	1.52	-3.9	-4.1	-0.9
19	265.1	-424.8	525.9	164.0	0.00	-0.5	3.0	-11.0	0.00	-3.5	-6.3	13.2
20	388.5	121.3	172.7	94.4	2.89	-0.5	3.9	-6.5	7.32	-3.1	-6.5	17.4
21	3090.6	343.9	258.2	2488.6	0.69	5.1	4.7	-1.2	1.04	-5.4	-5.9	1.4
Total	6920.4	-2500.4	2381.6	7039.0		0.9	4.6	-1.2		-5.6	-6.6	0.3

* Column (5) presents estimated coefficients after incorporating one intermediate observation. Columns (6-8) present %-differences between contributions obtained by the PB method incorporating one intermediate observation and obtained by calculating averages over traditional decomposition formulae (the latter are used as base values). Column (9) presents estimated coefficients after incorporating two intermediate observations. Columns (10-12) present %-differences between contributions obtained by the PB technique when incorporating two intermediate observations and obtained by the PB technique incorporating one intermediate observation (the latter are used as base values).

[a] Columns (2-4) do not always add up to the numbers in column (1) due to rounding.

The first column reports the actual change in sectoral labor costs in the 21 Spanish sectors between 1980 and 1994, expressed in billions of 1986 pesetas. Labor costs increased in all but two sectors, "energy" (2) and "other manufacturing" (13).

Columns (2)-(4) present the results of the decomposition analysis if the averages would have been taken over the six possible traditional decomposition formulae. The values for the Spanish economy as a whole in the bottom row are obtained by simply adding the sectoral results. Clearly, declining labor costs per unit of gross output would have led to lower labor costs in most sectors if nothing else would have changed (the Δ**u** effect is generally negative). Exceptions to this rule are mostly found in services sectors ("construction" (14), "communication services"(18), "real estate and business services" (20) and "other services" (21)). The generally positive results for the Δ**L** effect suggest that the domestic input coefficients have changed in such a way that labor costs would have increased, in the absence of changes in the level and composition of final demand. This finding can be due to changes in technology (technological progress or substitution of inputs induced by changes in relative prices) or to changes in the trade pattern of Spain. Only for "minerals and mining" (3) and "communication services" (18) a negative contribution of changes in the Leontief inverse **L** is obtained. Finally, the contributions of the $\Delta\mathbf{f}$ effect are positive. This is not a very surprising result, since the consumption and investment levels went up considerably in the period studied. In the demand-driven input-output model, positive final demand growth will always yield increases in labor costs, unless labor costs per unit of output are reduced considerably and/or input coefficients cause substitution of inputs towards inputs with lower labor cost coefficients. In SDA studies, these potentially offsetting effects are assumed to be absent by construction, however.

From the viewpoint of this paper, the most interesting results are contained in the righthand side part of Table 1. Column (5) presents the estimates for the θ_{f_i} parameters obtained by maximum entropy techniques for the PB method using the final demands for 1990 only.[22] The most prominent result concerns the frequency of extreme estimates. As often as five times a value of 0.00 was found, next to three times infinity. In these cases, the error committed by assuming a linear path may be very considerable, as polar paths appear more plausible.

Columns (6), (7) and (8) report the percentage deviations from the results obtained by computing averages as found after substituting the estimates in column (5) in equations (50) and (51). A couple of comments are called for. First, the deviations are sometimes considerable. The most substantial deviations are found for the $\Delta\mathbf{f}$ effect, associated with the determinant for which additional information was incorporated. Nevertheless, considerable effects of applying the PB methodology are also found for the other two effects. Second, the deviations are generally strongest for the sectors for which the estimated parameters deviate strongly from one, the value implicitly chosen when taking averages. The results for "transport services" (17), however, shows that deviations exceeding 10% can also occur for sectors for which the parameter deviates only modestly (1.71) from one. This property is due to the matricial nature of the decomposition at hand. Consequently, a different estimated time path for the final demand for commodities produced by sector *i* can well affect the contribution of changes in final demand as assessed for sector *j*. If the results are added over sectors, the most marked deviation from the decomposition using averages is found for the $\Delta\mathbf{L}$ effect. This results carries over from deviations in relative terms to deviations in its absolute value.

The results contained in column (9) are similar to those in (5), except that another observation for the final demand vector was incorporated in the estimation procedure. This clearly had impacts on the estimated θ_{f_i}. In five cases, an estimate below one in column (5)

22 The parameters relating to the determinants for which no additional information was used (the elements of the vector $\boldsymbol{\theta}_{\mathbf{u}}$ and the matrix $\boldsymbol{\Theta}_{\mathbf{L}}$) were set equal to one. As mentioned before, this implies that we assume a linear path for these determinants.

turned into an estimate above one after including the final demand levels for 1986. The most eye-catching change was found for sector 6, "metallic products", for which the estimated value is 6.71 now, as opposed to 0.30 before. Another interesting case is sector 14 "construction", for which the incorporation of more information yielded a switch from an extreme value for the parameter (0.00) to a value indicating an almost linear path (0.90). For the majority of sectors, however, the changes are relatively minor. The decomposition results in columns (10) to (12) show percentage differences between the respective contributions found with additional information for two years and for one year. With regard to the $\Delta\mathbf{u}$ effect, the most substantial changes are found for sectors 9 and 14, "transport equipment" and "construction", respectively. For these two sectors, the difference between the contributions of this effect exceeds 10 percent. Concerning the $\Delta\mathbf{L}$ effect, only one sector fulfills this admittedly rather arbitrary criterion for a change to be 'exceptional'. In sector 2, "energy", the difference between the contributions estimated using averages and the PB technique becomes much smaller if the latter method is applied with information not only for 1990, but also for 1986. With respect to the $\Delta\mathbf{f}$ effect, four sectors yield differences that exceed 10 percent: "other manufacturing" (13), "transport services" (17), "finance and insurance" (19) and "real estate and business services" (20). For the economy as whole, a notable result is that the inclusion of more additional information hardly affects the contribution of the effect of the determinant for which more information was included. The contribution of the other two effects changed more strongly. In both cases, the absolute size of the effects tend to be closer to zero than suggested by the use of contributions averaged over traditional decomposition formulae.

6.b Incorporating information from variables correlated to a determinant

As we explained in the previous section, additional information for other variables than the determinants can also be incorporated in the PB method, provided that this information concerns intermediate periods. We will illustrate this by using Spanish data on wages per unit of output (vector $\mathbf{r}$), a variable that is expected to be correlated quite strongly to the determinant $\mathbf{u}$, labor costs per unit of gross output. The data on $\mathbf{r}$ were taken from INE (1999). Our sample of 42 pairs (21 for 1980 and 1994 each) indicated a high correlation coefficient of 0.976.[23] We constructed 95%-confidence intervals for the elements of $\mathbf{u}$ in 1986 (equation (46)), using information on the wage costs per unit of gross output for that year taken from INE (1987). Next, we used the upper and lower bounds of these intervals to generate upper and lower bounds for the estimated parameters θ_{u_k}. The parameters for which no additional information was included (the $\mathbf{\Theta_L}$ matrix and $\boldsymbol{\theta}_\mathbf{f}$ vector) were all set equal to 1. The results are presented in Table 2.

[23] The estimated equation reads as follows: $u_k(t) = 0.030 + 1.099 r_k(t)$. The *t*-statistics are 3.48 and 28.72 for the interecpt and the slope parameters, respectively. The standard error of the equation is 0.026 and R^2 is 0.95.

Table 2: Decomposition results for case with additional information for a variable correlated to one of the determinants.*

Sector	(1) $\Delta\mathbf{c}$[a]	(2) $\Delta\mathbf{u}$ eff	(3) $\Delta\mathbf{L}$ eff	(4) $\Delta\mathbf{f}$ eff	(5) $\hat{\theta}_{u_k}^{low}$	(6) $\hat{\theta}_{u_k}^{high}$	(7) $\Delta\mathbf{u}$ eff low	(8) $\Delta\mathbf{u}$ eff high	(9) $\Delta\mathbf{L}$ eff low	(10) $\Delta\mathbf{L}$ eff high	(11) $\Delta\mathbf{f}$ eff low	(12) $\Delta\mathbf{f}$ eff high
1	24.2	-193.0	105.9	111.4	0.19	0.69	2.8	-15.5	-14.3	3.8	-13.3	1.3
2	-44.7	-132.6	35.2	52.7	0.92	∞	-5.3	-5.0	0.3	-6.0	-9.3	-12.7
3	53.2	-45.8	-76.0	175.0	0.00	0.00	-7.2	-11.9	-2.1	-2.8	-4.6	-2.8
4	42.5	-167.7	56.8	153.5	0.00	0.00	1.7	-3.4	-7.1	2.6	-1.1	0.9
5	261.5	-163.5	104.1	320.9	0.00	0.00	-21.6	-29.6	-4.6	3.0	-13.6	-12.0
6	130.0	-80.1	41.5	168.6	0.00	0.00	-12.4	-4.3	-9.1	-2.9	-3.6	-1.3
7	26.5	-122.5	49.8	99.2	0.00	0.00	0.3	-8.6	-8.0	3.2	-6.6	-1.2
8	216.6	-267.8	72.1	412.4	0.00	0.08	3.5	-34.5	-3.8	5.2	-22.2	1.4
9	232.0	-403.8	32.9	602.9	0.50	0.75	-6.9	-16.7	-15.3	-8.3	-10.3	-4.2
10	228.4	-15.5	98.6	145.3	0.00	1.03	3.0	-11.7	-0.4	-0.4	-0.9	0.6
11	249.5	-141.3	58.4	332.4	0.00	0.32	-9.9	-22.3	2.6	2.6	-9.9	-4.7
12	108.2	-74.4	86.5	96.1	0.00	0.00	-4.0	-15.7	-4.9	2.6	-7.7	-5.5
13	-77.9	-156.7	39.4	39.4	0.01	0.22	-2.0	2.4	-3.1	-9.0	10.8	6.9
14	570.1	9.7	53.4	506.9	∞	∞	17.1	17.1	-4.0	-4.0	0.1	0.1
15	527.0	-169.0	235.9	460.2	0.00	0.00	-9.6	-15.8	-3.9	0.5	-4.9	-3.8
16	163.6	-36.7	35.7	164.6	0.00	∞	16.0	-15.3	0.9	0.9	-2.6	3.4
17	359.7	-463.5	506.6	316.6	0.00	0.02	-10.7	-25.5	-11.2	0.8	-19.3	-15.9
18	105.8	83.5	-111.8	134.1	0.00	0.08	-1.1	1.4	8.8	6.0	4.1	8.1
19	265.1	-424.8	525.9	164.0	0.00	0.02	-1.4	-16.4	-12.3	-2.7	-4.6	3.4
20	388.5	121.3	172.7	94.4	0.00	0.76	-7.1	5.7	-3.3	3.5	-1.3	2.9
21	3090.6	343.9	258.2	2488.6	∞	∞	25.8	25.8	-4.2	-4.2	-4.4	-4.4
Total	6920.4	-2500.4	2381.6	7039.0			-9.3	-25.2	-8.4	-0.5	-6.8	-3.4

[*] Columns (1-5) are identical to columns (1-5) in Table 1, but are presented again for ease of reference. Columns (6-12) present %-differences between contributions obtained by the PB method and contributions obtained by calculating averages over traditional decomposition formulae (the latter are used as base values).

[a] Columns (2-4) do not always add up to the numbers in column (1) due to rounding.

The lower and upper estimates for the parameters that characterize the time paths indicate that extreme paths were found very often. For a number of sectors, one of the boundaries of the confidence interval indicates an extreme path, while the opposite boundary suggests a path that is strictly inside the rectangle. In general, the parameter estimates are very small. The zero estimates outnumber the infinity estimates, and the remaining estimates are generally below 1. Due to the strong correlation between the determinant **u** and the variable **r** for which we have information, the confidence interval for the path parameters turn out to be narrow. The most remarkable exception is sector 16, "restaurants, hotels, etc.", for which the additional information does not narrow the confidence interval at all.

The implications for the contributions of the respective determinants of using this kind of additional information in the PB approach are documented in columns (7-12) of Table 2. These columns indicate the relative differences between these results and those obtained by taking averages over traditional decomposition formulae, expressed in percentages. Clearly, the differences are sometimes sizeable. At the level of single sectors, differences of more than 20%

are encountered. The differences between the results obtained by using the set of lower boundaries and the results obtained by using the upper boundaries is often quite small, but not always. Such substantial differences are especially prominent in cases where one of the bounds characterizes a polar path, while the other does not. Examples are sectors 8 ("office equipment"), 11 ("textiles and clothing") and 17 ("transport services"). Another interesting feature concerns sectors for which the lower and upper bounds are identical to each other. In some cases, these extremely narrow confidence intervals yield identical contributions for the lower and upper bound values. Examples are sectors 14 ("construction") and 21 ("other services"). In other cases, such as sectors 5 ("chemicals") and 6 ("metallic products"), the results deviate significantly. The differences between these two sets of cases are due to the matricial nature of the decomposition at hand. The contributions depend on all θ parameters and changes in the determinants of all sectors. In most empirical input-output studies, most sectoral effects are mainly caused by intrasectoral changes, unless the sector considered is relatively small. Hence, it is not very surprising that equality of parameters for the two largest sectors in terms of the change in labor costs (sectors 14 and 21) implies equality of the contributions for these sectors, while stability of parameters for much smaller sectors like (5) and (6) does not preclude substantial differences in the relative sizes of the contributions for these sectors.

It should be kept in mind that we only present results for the sets of parameters that contain lower bounds for all sectors and upper bounds for all sectors. Of course, nothing prevents true parameter values to be close to the lower bounds for some sectors and close to the upper bounds for others. It might well be that some contributions would not be bounded by the levels implied by the results presented in Table 2. More advanced statistical analysis is required to study this issue in depth.

In this section, we analyzed empirically how the incorporation of two specific kinds of additional information into the PB approach affect results for a sectoral labor costs SDA for Spain in the period 1980-1994. We would like to stress that we could also have opted for incorporation of the two types of information simultaneously. This would not involve a more complex decomposition methodology. The only reason we did not embark on such an endeavor relates to limited space.

7. Conclusions

Traditional SDA suffers from the "non-uniqueness" problem. Since many decomposition formulae are equally valid from a theoretical point of view, the empirically often substantial differences in outcomes noted by Dietzenbacher & Los (1998) pose a serious problem. Most often, the problem of choosing a specific formulae is avoided by computing averages over (a subset of) formulae. This paper does not challenge the theoretical equivalence of decomposition formulae but proposes a methodology using Maximum Entropy econometrics to select the decomposition formula that provides an optimal 'fit' to additional empirical information.

The point of departure is a class of monotonic time paths for variables, which led us to label our method the "path based" (PB) method. It was shown that taking the average over all traditional decomposition formulae is equivalent to one specific member of this class, i.e. the linear path. Next, we showed how the parameters that characterize the paths can be estimated, even if the available data is very limited. If information about the values of the determinants contained in the SDA is completely absent, the estimation procedure yields the linear path. If some information is available for a period between the initial period and the final period of the analysis, the selected path is a different one. In some cases, the selected path corresponds to paths that are implicitly assumed by specific traditional decomposition formulae. Together, the estimated parameters define a decomposition formula. From an empirical point of view, this formula is to be preferred over other decomposition formulae that can be constructed by means of the monotonic times paths considered.

We applied the methodology to quantify the contributions of three determinants of changes in sectoral labor costs in Spain between 1980 and 1994, i.e. labor costs per unit of gross output, input coefficients and final demand levels. Two types of additional information were considered. First, the actual levels of final demand for one or two intermediate years, and second, the levels of sectoral wage costs per unit of output in an intermediate period. These wage costs are used as a variable that correlates with one of the determinants, the labor costs per unit of gross output. The results indicate that the use of additional information in the PB approach can well yield results that differ substantially from the mean over all traditional decomposition formulae, or equivalently, the linear path. For some sectors and determinants, the differences amount to more than 10%. Differences of this size lead us to believe that the PB method provides an interesting alternative to computing averages over decomposition formulae.

A couple of challenges remain to be solved, however. In a considerable number of cases, the additional information did not fit the class of monotonic paths we defined. We opted for a rather pragmatic solution if the value of a determinant in an intermediate period exceeded the values in both the initial and the final period (or if it was lower than both), which implies non-monotonicity. We would of course prefer an approach in which non-monotone paths could be estimated. More research should be done in this respect, because a more general class of time paths would complicate the construction of the constrained maximization problems characteristic of maximum entropy estimation procedures. It could also be interesting to see whether estimation results would change if we would estimate the parameters in a way that takes the continuous time character of the temporal paths explicitly into account. In this paper, we do implicitly assume that final demand levels are constant over a year, which is not really in line with the continuous nature of the temporal paths considered. This seems to be a very ambitious task, however.

Another challenge is the application of the PB principle to another type of structural decomposition analyses. In this paper, we only considered additive decomposition forms, which quantify contributions of changes in determinants to *differences* in values between final and initial periods. Recently, multiplicative decomposition forms have become increasingly popular. They quantify contributions of changes in determinants to *ratios* in values between final and initial periods. Despite the difference between the two types of decomposition, they both suffer from the non-uniqueness problem. Hence, it could be worthwhile to pursue a PB alternative to the (geometric) averages over multiplicative decomposition formulae that are mostly used to circumvent the problem.

References

Barff, R..A. and P.L. Knight III (1988), "Dynamic Shift-Share Analysis", *Growth and Change*, vol. 19, pp. 2-10.

De Haan, M. (2001), "A Structural Decomposition Analysis of Pollution in The Netherlands", *Economic Systems Research*, vol. 13, pp. 181-196.

Dietzenbacher, E., A.R. Hoen and B. Los (2000), "Labor Productivity in Western Europe 1975-1985: An Intercountry, Interindustry Analysis", *Journal of Regional Science*, vol. 40, pp. 425-452.

Dietzenbacher, E., M.L. Lahr and B. Los (2004), "The Decline in Labor Compensation's Share of GDP: a Structural Decomposition Analysis for the United States, 1982-1997", in: E. Dietzenbacher and M.L. Lahr (eds.), *Wassily Leontief and Input-Output Economics* (Cambridge UK: Cambridge University Press), pp. 138-185.

Dietzenbacher, E. and B. Los (1998), "Structural Decomposition Techniques: Sense and Sensitivity", *Economic Systems Research*, vol. 10, pp. 307-323.

Dietzenbacher, E. and B. Los (2000), "Structural Decomposition Analyses with Dependent Determinants", *Economic Systems Research*, vol. 12, pp. 497-514.

Fernández, E. (2004), *The Use of Entropy Econometrics in Decomposing Structural Change*, unpublished PhD thesis, University of Oviedo (Spain).

Golan, A., G. Judge and D. Miller (1996), *Maximum Entropy Econometrics: Robust Estimation with Limited Data* (Chichester UK, John Wiley).

Golan, A., G. Judge and S. Robinson (1994), "Recovering Information from Incomplete or Partial Multisectoral Economic Data", *Review of Economics and Statistics*, vol. 76, pp. 541-549.

Harrison, W.J., J.M. Horridge and K.R. Pearson (2000), "Decomposing Simulation Results with Respect to Exogenous Shocks", *Computational Economics*, vol. 15, pp. 227-249.

Hoekstra, R. and J.C.J.M. Van den Bergh (2002), "Structural Decomposition Analysis of Physical Flows in the Economy", *Environmental and Resource Economics*, vol. 23, pp. 357-378.

INE (1987), *Contabilidad Nacional de España 1986* (Madrid, Instituto Nacional de Estadística).

INE (1991), *Contabilidad Nacional de España 1990* (Madrid, Instituto Nacional de Estadística).

INE (1998), *Tabla input-output de España 1994* (Madrid, Instituto Nacional de Estadística).

INE (1999), *Contabilidad Nacional. Serie enlazada 1964-1991.* (Madrid, Instituto Nacional de Estadística).

Kapur, J.N. and H.K. Kesavan (1993), *Entropy Optimization Principles with Applications* (New York: Academic Press).

Robinson, S., A. Cattaneo and M. El-Said (2001), "Updating and Estimating a Social Accounting Matrix Using Cross-Entropy Methods", *Economic Systems Research*, vol. 13, pp. 47-64.
Rose, A. and S. Casler (1996), "Input-Output Structural Decomposition Analysis: A Critical Appraisal", *Economic Systems Research*, vol. 8, pp. 33-62.
Shannon, J. (1948), "A Mathematical Theory of Communication", *Bell System Technical Bulletin Journal*, vol. 27, pp. 379-423.
Sun, J.W. (1998), "Changes in Energy Consumption and Energy Intensity: A Complete Decomposition Model", *Energy Economics*, vol. 20, pp. 85-100.

Appendix A: Construction of Data for Empirical Illustration

This appendix briefly depicts the main sources of information consulted for the empirical study in Section 6, as well as the manipulations we had to carry out before we could apply the decomposition analysis to changes in sectoral labor costs in Spain. The original tables for 1980 and 1994 were inconvenient in that they were not directly comparable in several respects, such as their sectoral classification, the prices they were expressed in and the methodology used for their construction. These features will be explained in more detail below.

The first source of the lack of homogeneity in these two tables is due to the fact that the original 1980 table does not include Value Added Tax (VAT), while the table for 1994 does. The Spanish Statistical Institute (INE) published a series of harmonized input-output tables for 1964 to 1991 (INE, 1999) that were homogenized in such a way that they all include VAT. We used the input-output table of 1980 as well as the final demands of 1986 and 1990 (as additional information in Section 6) from this source.

The second problem concerned the different sectoral classifications of the input-output tables. The 1980 table specified 30 sectors, whereas in the 1994 table as many as 57 sectors were distinguished. Aggregation led to a common classification that discerns 21 sectors (see Table A.1), for which price deflators were available (see below). Especially with respect to services sectors, we had to aggregate rather rigorously.[24]

24 Details can be found in Fernández (2004, pp. 54-60).

Table A.1: Sector classification and aggregation scheme

Sector	Name	Sectors in original table for 1980	Sectors in original table for 1994
1	Agriculture	01	1
2	Energy	06	2+3+4+5+6+7+8+9+10+11
3	Minerals and mining products	13	12+13
4	Non-metallic products	15	14+15+16+17
5	Chemical products	17	18
6	Metallic products excepting transport equipment	19	19
7	Machinery for agriculture and industry	21	20
8	Office equipment, measuring equipment and others	22+25	21+22
9	Transport equipment	28	23+24
10	Food, drinks and tobacco	36	25+26+27+28+29
11	Textiles, leather and clothing	42	30+31
12	Paper and derived products	47	33+34
13	Industries not elsewhere classified	48+49	32+35+36
14	Building materials and construction	53	37
15	Commerce an repairing services	56	38+39
16	Restaurants, hotels and cafes	59	40
17	Transport services	611+613+63+67	41+42+43+44+45
18	Communications	69	46
19	Finance and insurance	71	47+48
20	Real estate and services to companies	73+75	49+50
21	Other services	77+79+86	51+52+53+54+55+56+57

The final source of incomparability is related to prices. As a matter of fact, the tables were expressed in current prices. Since a good intertemporal comparison required us to remove any distortions caused by changes in relative price levels rather than changes in quantities of required inputs, we had to apply a deflation procedure. A series of implicit deflators has been computed employing data from the information of sectoral Gross Values Added at market prices, available for several years (INE, 1999). The series were published expressed in current prices ($\mathbf{gva}_{\text{cump}}$) as well as in constant prices of 1986 ($\mathbf{gva}_{\text{comp}}$). This information was used to obtain a deflator for every sector, although it is an admittedly crude procedure. The mathematical expression for these implicit deflators (for sector *i*) is

$$d_{t,86}^{i} = \frac{gva_{cump,i}}{gva_{comp,i}} \tag{A.1}$$

Therefore, the variables of the input-output tables could be expressed in 1986 prices employing $d^{i}_{t,86}$ as deflator from the expression

$$q^{d}_{ij} = \frac{q^{t}_{ij}}{d^{i}_{t,86}} \qquad \text{(A.2)}$$

where q^{d}_{ij} can denote intermediate deliveries, labor costs or final demand in 1986 prices and q^{t}_{ij} is this same flow in current prices. This procedure is applied to the values in the 1994 input-output table (INE, 1998), as well as to the 1990 final demands (INE, 1991) used as a source of additional information in Section 6. We did not need to apply the same to the 1980 input-output tables, since the Spanish national statistical agency published this table already in 1986 prices (INE, 1999). Taking as reference an intermediate period like 1986 implies that the adjustments in the prices are smoother than if 1980 or 1994 had been taken as the base instead.