Probabilistic Sensitivity Analysis in Input-Output Models

Marco Percoco*
Department of Economics
Bocconi University

June 25, 2004

Abstract

Following the seminal work by Bullard and Sebald [Effects of Parametric Uncertainty and Technological Change on Input-Output Models, *Rev. of Ec. and Stat.*, vol. 59, 75-81], in this paper we present an innovative approach to sensitivity analysis in Input-Output model. In particular, we propose a statistical model capable to compute a sensitivity index associated to each technical coefficient. We call the ordered set of these indices Importance Matrix. Finnally, in order to show a simple example for this methodology, we consider the case of the Chicago economy.

Keywords: Input-Output Models, Sensitivity Analysis, Importance Matrix

JEL Classification: C15, C67, D5

*This paper was written while the author was visiting the Regional Economics Applications Lab., University of Illinois at Urbana-Champaign. I would like to thank Stefano Tarantola and the Joint Research Center of the European Union for providing the SimLab software, Geoffrey Hewings for valuable comments and for providing the data and Chokri Dridi for stimulating discussions. Usual disclaimer applies. *Correspondence to:* Marco Percoco, Istituto di Economia Politica "E. Bocconi", Università Bocconi, via Gobbi 5, 20136 Milano, ITALY. e-mail: marco.percoco@uni-bocconi.it.

1 Introduction

One of the main issues in input-output (I-O) models is the fact that they are substantially deterministic in the technological parameters. But, as pointed out in several studies (Bullard and Sebald, 1977; Jackson, 1986; Israilevich *et alii*, 1996), there are at least four sources of uncertainty in I-O models:

a) if the I-O model is survey-based, then there could be a classical sampling error;

b) in the case of large surveys, an error in the inference design can raise;

c) real technical coefficients are not constant over the time, and in an age of structural change due to technological development, this error can be relevant;

d) errors in compiling the large data base can affect the quality of the final table.

These issues have been addressed by scholars since the beginnings of the theory on I-O analysis. Quandt (1958 and 1959), for example, runs a primitive sensitivity analysis by disturbing the error distributions and then observing the change in output.

Since then, three different paradigms, attempting to analyze the stochastic behavior of technical coefficients in an I-O framework, have raised. First, following the seminal work of Simonovitz (1975), Lahiri (1983) and Lahiri and Satchel (1985) have provided some conditions for the over- and underestimation of the Leontief matrix (and of the output) by assuming biproportional stochastic independence in the elements.

The second approach relies on the theoretical work by Sherman and Morrison (1950) who analyzed the effect on the Leontief inverse of a change in an element of the original matrix. In this context, Sonis and Hewings (1992) developed the well-known theory of the fields of influence and of the deterministic sensitivity analysis. On the other hand, West (1986) extended the results of Sherman and Morrison (1950) by considering the case of a stochastic Leontief inverse.

The third paradigm is hybrid, in the sense that it considers the possibility of updating the I-O tables through econometric models (see Kraybill, 1991; Conway, 1990; Treyz and Stevens, 1985; Treyz, 1993). In this context, Israilevich *et alii* (1997) provide an interesting approach to structural change forecasting by considering an innovative framework based on the general assumptions of the computable general equilibrium models.

The present paper tries to increase the volume of a partially abandoned approach of the economic literature, that is the simulation analysis in the context of I-O models. In particular, since the seminal work by Bullard and Sebald (1988) appeared, few papers have provided further insights. But, as stated in Jackson (1986), there are significative reasons to consider a probabilistic approach in an I-O context. In particular, he writes that *"the set of all like coefficients for an industry for each of m regions defines m subpopulations, and an associated probability of realizing a particular coefficient within the total population or within each subpopulation"*. In addition, as long as we consider a probabilistic approach, the term "error" is no longer appropriate because the probability distribution associated to each technical coefficient is meant to provide the complete range of possible realizations.

The aim of this article is to present an innovative procedure in order to run quantitative sensitivity analysis in the context of I-O models. In particular, we propose the computation of an *importance matrix* defined as the ordered set of the indices of sensitivity associated to each technical coefficient of the Leontief matrix. This new concept provides a quantitative measure of the relative importance for the economy of each element and of each sector.

The paper is organized as follows. In Section 2 we present the statistical procedure in order to compute the sensitivity indices through simulation; in Section 3 a simple example based upon the Chicago I-O Table is illustrated. Finally, in Section 4 we provide some concluding remarks and propose some paths of future research.

2 The Statistical Model

Let us consider the classical I-O model in the form:

$$\mathbf{Y}=(\mathbf{I}\text{-}\mathbf{A})^{-1}\mathbf{X} \tag{1}$$

where $\mathbf{Y}$ is the vector $n \times 1$ of the total output of the n sectors in the economy, $\mathbf{X}$ is the vector $n \times 1$ of the final demand and $(\mathbf{I}\text{-}\mathbf{A})^{-1}$ is the classical inverse matrix of the technical coefficients of dimension $n \times n$. In general, we can write (1) as:

$$\mathbf{Y}{=}f\left[\mathbf{(I\text{-}A)}^{-1},\mathbf{X}\right] \tag{2}$$

The aim of this paper is to assess the volatility in the elements of $\mathbf{(I\text{-}A)}^{-1}$ and how it explains the variance V of the total output. Let us suppose that all these coefficients are affected by uncertainty and that the elements of $\mathbf{X}$ are known, so that, for the purpose of Sensitivity Analysis (SA) we can write:

$$\mathbf{Y}{=}f\left(a_{1i},...,a_{ij},...,a_{nn}\right)$$

with $a_{ij}=[a]_{ij}\in\mathbf{(I\text{-}A)}^{-1}$. If the generic element a_{ij} is fixed to a generic value $\widetilde{a}_{ij}$, then the variance of sector y_i can be written as:

$$\begin{aligned} V\left(y\left|a_{ij}=\widetilde{a}_{ij}\right.\right) &= \int...\int\left[f\left(a_{i1},...,\widetilde{a}_{ij},...,a_{in}\right)-E\left(y_i\left|a_{ij}=\widetilde{a}_{ij}\right.\right)\right]^2\prod_{i\neq j}p_i(a_i)\mathrm{d}a_i = \\ &= \int...\int f\left(a_{i1},...,\widetilde{a}_{ij},...,a_{in}\right)\prod_{i\neq j}p_i(a_i)\mathrm{d}a_i-\left[E\left(y_i\left|a_{ij}=\widetilde{a}_{ij}\right.\right)\right]^2 \end{aligned} \tag{3}$$

For the purpose of sensitivity analysis one is interested in eliminating the dependence upon the value $\widetilde{a}_{ij}$ by integrating $V\left(y\left|a_{ij}=\widetilde{a}_{ij}\right.\right)$ over the probability density function of $\widetilde{a}_{ij}$, obtaining:

$$E\left[V\left(y\left|a_{ij}\right.\right)\right]=\int...\int\left[f\left(a_{i1},...,a_{ij},...,a_{in}\right)\right]^2\prod_{i}p_i(a_i)\mathrm{d}a_i-\int\left[E\left(y_i\left|a_{ij}=\widetilde{a}_{ij}\right.\right)\right]p_j(\widetilde{a}_{ij})\mathrm{d}\widetilde{a}_{ij} \tag{4}$$

Notice that we have dropped the dependence $\widetilde{a}_{ij}$ from the left-hand side, as it disappears due to the integration.

Let us define the variance of y_i as

$$V(y_i)=\int...\int\left[f\left(a_{i1},...,a_{ij},...,a_{in}\right)\right]^2\prod_{i}p_i(a_i)\mathrm{d}a_i-\left[E\left(y_i\right)\right]^2 \tag{5}$$

By subtracting Eq. (4) from Eq. (5) we obtain:

$$V(y_i)-E\left[V\left(y_i\left|a_{ij}\right.\right)\right]=\int\left[E\left(y_i\left|a_{ij}=\widetilde{a}_{ij}\right.\right)\right]^2p_j(\widetilde{a}_{ij})\mathrm{d}\widetilde{a}_{ij}-\left[E\left(y_i\right)\right]^2 \tag{6}$$

By definition, we have that $V(y_i)-E\left[V\left(y_i\left|a_{ij}\right.\right)\right]=V\left[E\left(y_i\left|a_{ij}\right.\right)\right]$, and it is a good measure of the sensitivity of y_i with respect to the technical coefficient a_{ij}. If we divide it by the unconditional variance, we obtain the *index of sensitivity* :

$$S_{ij} = \frac{V\left[E\left(y_i \left| a_{ij}\right.\right)\right]}{V(y_i)} \tag{7}$$

that is scaled in [0,1]. The problem is that Eq. (6) is computationally impractical. In a Monte Carlo (MC) framework, it implies a double loop: the inner to compute $\left[E\left(y_i\right)\right]^2$ and outer to compute the integral. For this reason, we can rewrite the (6) as (Ishigami and Homma, 1990):

$$\begin{aligned} V(y_i) - E\left[V\left(y_i \left| a_{ij}\right.\right)\right] &= \int ... \int f\left(a_{i1}, ..., a_{ij}, ..., a_{in}\right) f\left(a'_{i1}, ..., a'_{ij}, ..., a'_{in}\right) \times \\ &\times \prod_i p_i(a_i)\mathrm{d}a_i \prod_i p_i(a'_i)\mathrm{d}a'_i - \left[E\left(y_i\right)\right]^2 \end{aligned} \tag{8}$$

The expedient of using the additional integration variable primed, allow us to realize that the integral in the previous equation is the expectation value of the function f on a set of $(2n-1)$ technical coefficients. Now, this integral can be computed using a single MC loop, as argued by Saltelli et al. (1993).

Let us generate two sample matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ for the technical coefficients:

$$\mathbf{M}_1 = \begin{pmatrix} a_{111} & a_{112} & ... & a_{11n} & ... & a_{1ij} & ... & a_{1nn} \\ a_{211} & a_{212} & ... & a_{21n} & ... & a_{2ij} & ... & a_{2nn} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ a_{k11} & a_{k12} & ... & a_{k1n} & ... & a_{kij} & ... & a_{knn} \end{pmatrix}$$

$$\mathbf{M}_2 = \begin{pmatrix} a'_{111} & a'_{112} & ... & a'_{11n} & ... & a'_{1ij} & ... & a'_{1nn} \\ a'_{211} & a'_{212} & ... & a'_{21n} & ... & a'_{2ij} & ... & a'_{2nn} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ a'_{k11} & a'_{k12} & ... & a'_{k1n} & ... & a'_{kij} & ... & a'_{knn} \end{pmatrix}$$

where k is the sample size used for the MC estimate. In order to estimate the sensitivity measure for the generic technical coefficient a_{ij}, i.e.

$$S_{ij} = \frac{V\left[E\left(y_i \left| a_{ij}\right.\right)\right]}{V(y_i)} = \frac{U_{ij} - \left[E\left(y_i\right)\right]^2}{V(y_i)}$$

$$U_{ij} = \int \left[E\left(y_i \left| a_{ij} = \widetilde{a}_{ij}\right.\right)\right]^2 p_j(\widetilde{a}_{ij}) \mathrm{d}\widetilde{a}_{ij}$$

we need an estimate for both $E(y_i)$ and U_{ij}. The former can be obtained from the values of y_i computed on the sample in $\mathbf{M}_1$ or $\mathbf{M}_2$. U_{ij} can be obtained from values of y_i computed on the following matrix $\mathbf{N}_{ij}$:

$$\mathbf{N}_{ij} = \begin{pmatrix} a'_{111} & a'_{112} & ... & a'_{11n} & ... & a_{1ij} & ... & a'_{1nn} \\ a'_{211} & a'_{212} & ... & a'_{21n} & ... & a_{2ij} & ... & a'_{2nn} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ a'_{k11} & a'_{k12} & ... & a'_{k1n} & ... & a_{kij} & ... & a'_{knn} \end{pmatrix}$$

i.e. by

$$\widehat{U}_{ij} = \frac{1}{k-1} \sum_{r=1}^{k} f(a_{r1}, ..., a_{rn}) f(a'_{r1}, ..., a'_{rn}) \tag{9}$$

If one thinks of matrix $\mathbf{M}_1$ as the "sample" matrix, and of $\mathbf{M}_2$ as the "re-sample" matrix, the $\widehat{U}_{ij}$ is obtained from products of values of f computed from the sample matrix times values of f computed from $\mathbf{N}_{ij}$, i.e. a matrix where all factors except a_{ij} are re-sampled. In this way the computational cost associated with a full set of first order indices S_{ij} (with $i = 1, ..., n$ and $j = 1, ...n$) is $k(n+1)$. One set of k evaluations of f is necessary to compute $E(y_i)$, and n sets of k evaluations of f are needed for the second term of the product in Eq. (9). This means that the total cost of this procedure is $2k(n+1)$.

It could be interesting to note that the complete set of sensitivity indices in Eq. (7) is a $n \times n$ matrix and it provides a more specific formulation of the importance measures provided by Bullard and Sebald (1977). In what follows, this set will be called *Importance Matrix.*

3 Experiment Design and Results

In order to provide a simple example of the procedure described in Section 2, we use the Chicago I-O table for the year 2000[1]. This table contain just 9 sectors (i.e. 81 parameters): Resources (RES), Construction

[1]See Israilevich et al. (1997) for further details

(CONST), Non-Durable Goods (NDG), Durable Goods (DG), Transportation (TRANS), Trade (TRADE), Financial Services and Real Estate (FIRE), Services (SER) and Government (GOV).

The first step is to assign a distribution function to the technical coefficients. According to Bullard and Sebald (1988), but partially different from Jackson (1987), we will consider a lognormal distribution with a 99.7% confidence interval. In addition, in order to account for the dependence among the parameters, we compute the correlation coefficients using the time series of the I-O tables over the period 1975-2000[2].

After 1000 runs of the 9 equations of the model, the first result is shown in Figure 1[3].

(Figure 1, Page 14)

As expected, the distribution of the sectorial outputs approximates the lognormal.

In Table 1 we present the *Importance Matrix* $\mathbf{S}_{(n\times n)}$. On the columns there are the reacting sector, i.e. the sectors $i = 1, ..., n$ affected by a change in the technical coefficient a_{ij}. In the rows there are the activating sector, i.e. the sectors $j = 1, ..., n$ whose technological change is meant to generate uncertainty on the sectors $i = 1, ..., n$. This, in turn, means that the generic element of the matrix $S_{ij} = [s]_{ij} \in\mathbf{S}$ measures the effect on the output of the sector i of a change of the technical coefficient a_{ij}.

As expected, for the Resources and Construction sectors, the variables SER, TRADE and DG present the highest values (Figg. 2 and 3). In addition, the transportation sector (TRANS) has a great importance for the durable goods industry (Fig. 5), whilst the public expenditure does not show relevant values (Figg. 2-10), only for the RES sector it shows a Pearson Coefficient somehow different from zero.

[2]For space-constraints, the complete correlation matrix is omitted (there are 6561 elements); it may be obtained on request to the author in E.views format.

[3]We use the SimLab software, kindly provided by the Joint Research Center of the European Union, Ispra, Italy. In addition we use the explicit formulation of the technical coefficients so that we do not need to apply an exclusion procedure for the sample in which the column sum differs from 1. It should also noted that by using an assumption on the probability function and on the confidence interval, the structure of the simulated matricis are consistent with the theory of entropy in I-O systems.

(Figure 2-10, Pages 14 to 18)

In the context of the importance matrix, we can compute the following marginal index:

$$S_i = \sum_j S_{ij} \tag{10}$$

which is meant to measure the absolute importance of the sector j for the economy as a whole (we will call it *index of absolute importance*). Note that $S_i : [0,1] \to [0,n]$. In that case, as shown in Table 2, durable goods industry and services present the highest values, implying a strong dependence of the Chicago economy from these two sectors.

(Table 2, Page 12)

In addition, we can compute a synthetic index of reaction by considering:

$$S_j = \sum_i S_{ij} \tag{11}$$

This index (we will call it *index of absolute sensitivity;* $S_j : [0,1] \to [o,n]$) provides a measure of the aggregate volatility of the n sectors. Table 3 shows the results for the Chicago economy. Even if the difference among the indices of absolute sensitivity is quite narrow ($\sigma_{S_j} = 0.118$ and t-statistics $= 20.513$), durable goods sector again and RES present the highest values, implying an higher sensitivity of these sector to structural changes in the economy.

(Table 3, Page 13)

4 Conclusions

In this paper we have provided a new and promising methodology to assess the effect on the output of the uncertainty in the technical coefficients of an I-O model. After describing the simulation design and the computational procedure, we have proposed four innovative importance measures: theindex of sensitivity , the Importance Matrix, the index of absolute importance and the index of absolute sensitivity. These are meant to describe the structure of the effect on the sectors of the economy of a structural change .

We have applied this model to a simple 9 sectors I-O table for the Chicago economy for the year 2000. We have found an expected great importance of the services and a surprising stability (in terms of low indices of both importance and sensitivity) for the financial services, insurance and real estate. It could be a sign of substantial independence of the considered sector from the other urban industries.

Future research could address the comparison of this technique with the deterministic theory of fields of influence. In addition, the quantitative sensitivity analysis could be used in the context of updating/forecasting the technical parameters by providing a measure of relative importance for each a_{ij} and then finding out the ones that could be forecasted (with low S_{ij}) and the ones needing a survey design because more important for the economy.

References

[1] Bullard, C.W. and A. Sebald (1977), Effects of Parametric Uncertainty and Technological Change in Input-Output Models, *Review of Economics and Statistics*, vol. 59, No.. 1, 75-81

[2] Bullard, C.W. and A. Sebald (1988), Monte Carlo Sensitivity Analysis in Input-Output Models, *Review of Economics and Statistics*, vol. 70, No. 6, 708-712

[3] Conway, R.S. (1990), The Washington Projection and Simulation Model: Ten Years of Experience with a Regional Interindustry Econometric Model, *International Regional Science Review*, vol. 13, 141-165

[4] Cukier, R.I., C. Fortuin, C.M. Schuler, K.E. Petschek and J.H. Schaibly (1973), Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients, *Journal of Chemical Physics*, vol. 59, 3873-3878

[5] Ishigami, T. and T. Homma (1990), An Importance Quantification Technique in Uncertainty Analysis for Computer Models, in *Proceedings of the ISUMA '90, First International Symposium on Uncertainty Modeling and Analysis*, University of Maryland

[6] Israilevich, P.R., G.J.D. Hewings, M. Sonis, G.R. Schindler (1997), Forecasting Structural Change With A Regional Econometric Input-Output Model, *Journal of Regional Science*, Vol. 37, No. 4, 565-590

[7] Israilevich, P.R., G.J.D. Hewings, G.R. Schindler and R. Mahidhara (1996), The Choice of An Input-Output Table Embedded in Regional Econometric Input-Output Models, *Papers in Regional Science*, vol. 75, No. 2, 103-119

[8] Jackson, R.W. (1986), The Full-Distribution Approach to Aggregate Representation in the Input-Output Modeling Framework, *Journal of Regional Science*, vol. 26, No. 3, 515-531

[9] Kraybill, D.S. (1991), Multiregional Computable General Equilibrium Models: An Introduction and Survey, Dept. of Agricultural and Applied Economics, University of Georgia, mimeo

[10] Lahiri, S. (1983), A Note on the Underestimation and Overestimation in Stock Input-Output Models, *Economics Letters*, vol.13, 361-366

[11] Lahiri, S. and S. Satchell (1985), Underestimation and Overestimation of the Leontief Inverse Revisited, *Economics Letters*, vol. 18, 181-186

[12] Quandt, R. (1958), Probabilistic Errors in the Leontief Systems, *Naval Research Logistics Quarterly*, vol. 5, 155-170

[13] Quandt, R. (1959), On the Solution of Probabilistic Leontief Systems, *Naval Research Logistics Quarterly*, vol. 6, 295-305

[14] Saltelli, A., S. Tarantola and K. Chan (1999), A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output, *Technometrics*, vol. 41, No. 1, 39-56

[15] Saltelli, A., K. Chan and M. Scott (2000), *Sensitivity Analysis*, John Wiley and Sons, New York

[16] Saltelli, A., S. Tarantola and F. Campolongo (2000), Sensitivity Analysis as an Ingredient of Modeling, *Statistical Science*, vol. 15, No. 4, 377-395

[17] Saltelli, A. and S. Tarantola (2003), On the Relative Importance of Input Factors in Mathematical Models, *Journal of the American Statistical Association*, forthcoming

[18] Sherman, J. and W.J. Morrison, Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix, *Annals of Mathematical Statistics*, vol. 21, 124-127

[19] Simonovits, A. (1975), A Note on the Underestimation and Overestimation of the Leontief Inverse, *Econometrica*, 124-127

[20] Sonis, M. and G.J.D. Hewings (1992), Coefficient Change in Input-Output Models: Theory and Applications, *Economic Systems Research*, No. 4, 143-157

[21] Treyz, G.I. (1993), *Regional Economic Modeling*, Kluwer, Boston

[22] Treyz, G.I. and B.H. Stevens (1985), The TFS Modeling Methodology*, Regional Studies,* vol. 19, 547-562

[23] West, G.R. (1986), A Stochastic Analysis of an Input-Output Model, *Econometrica*, vol. 54, No. 2, 363-374

Table 1: Importance Matrix

	RES	CONST	NDG	DG	TRANS	TRADE	FIRE	SER	GOV
RES	0.010	0.085	0.009	0.037	0.102	0.074	0.025	0.079	0.063
CONST	0.231	0.126	0.206	0.253	0.194	0.284	0.158	0.223	0.119
NDG	0.287	0.224	0.437	0.273	0.322	0.335	0.336	0.288	0.261
DG	0.410	0.390	0.450	0.438	0.505	0.470	0.466	0.440	0.454
TRANS	0.276	0.296	0.161	0.370	0.294	0.272	0.127	0.263	0.291
TRADE	0.428	0.414	0.402	0.410	0.236	0.387	0.413	0.328	0.466
FIRE	0.165	0.085	0.054	0.157	0.114	0.045	0.002	0.034	0.088
SER	0.610	0.598	0.661	0.624	0.667	0.562	0.661	0.757	0.653
GOV	0.104	0.067	0.009	0.067	0.037	0.020	0.044	0.004	0.033

Table 2: Indices of Absolute Importance

Sector	S_i
SER	5.793
DG	4.023
TRADE	3.484
NDG	2.763
TRANS	2.35
CONST	1.794
FIRE	0.744
RES	0.484
GOV	0.3849

Table 3: Indices of Absolute Sensitivity

Sector	S_j
DG	2.629
RES	2.521
TRANS	2.471
TRADE	2.449
GOV	2.428
SER	2.416
NDG	2.389
CONST	2.285
FIRE	2.232

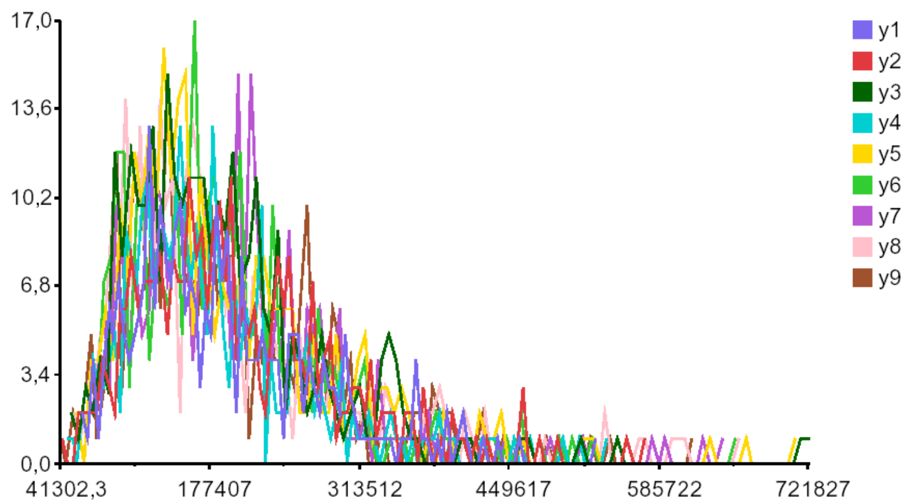

Figure 1: Patterns of Simulated Sector Outputs

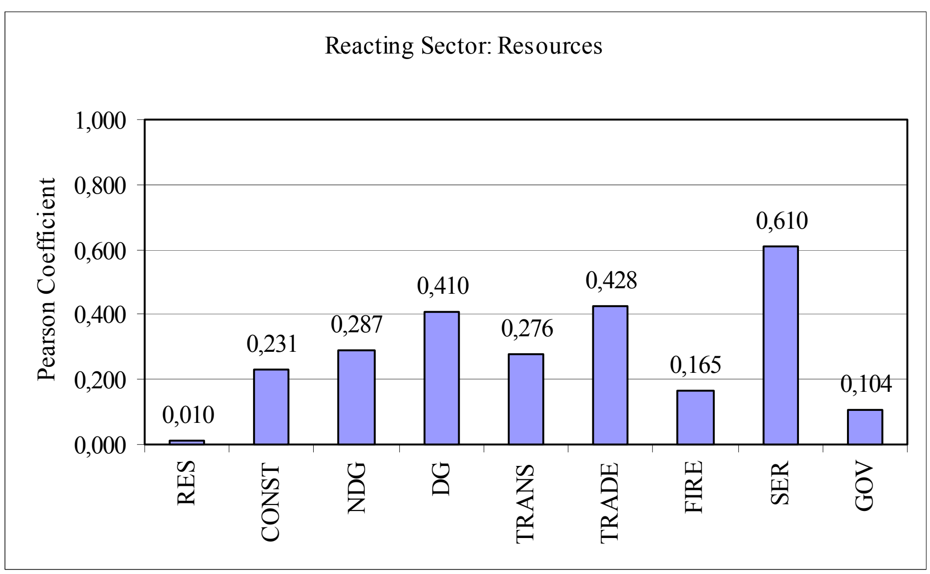

Figure 2:

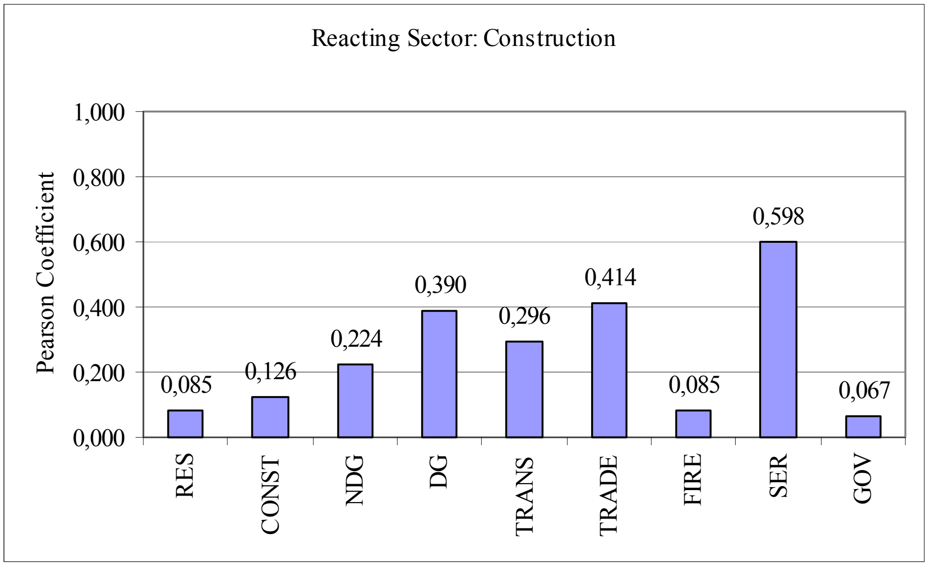

Figure 3:

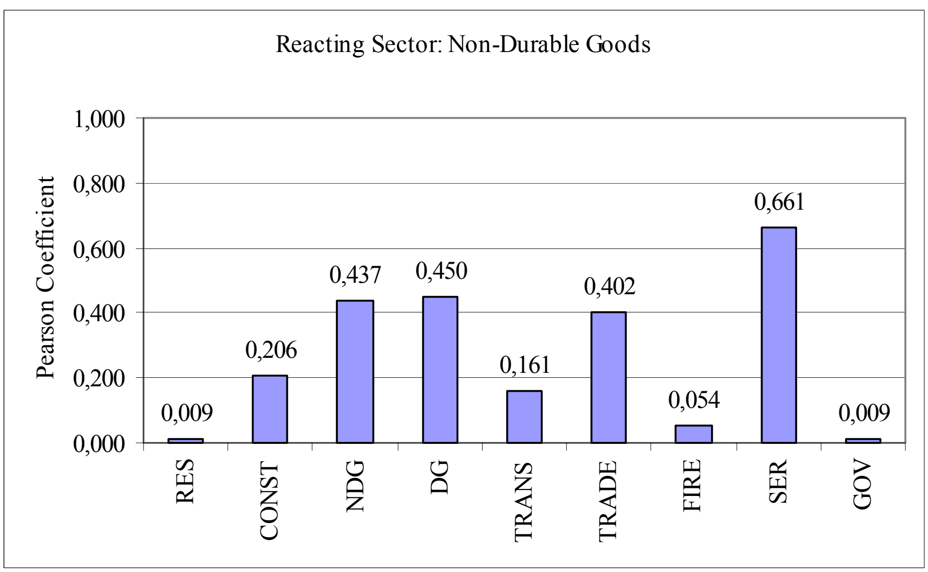

Figure 4:

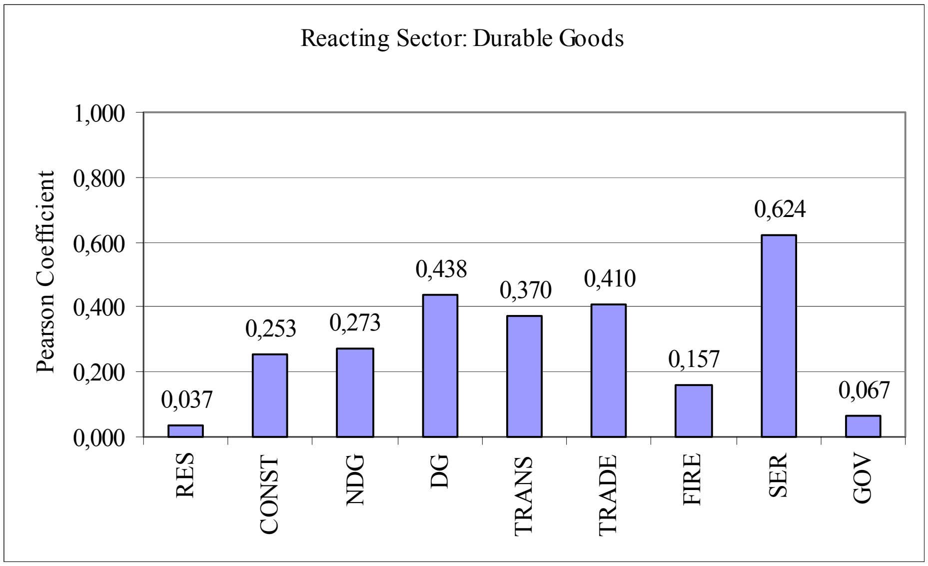

Figure 5:

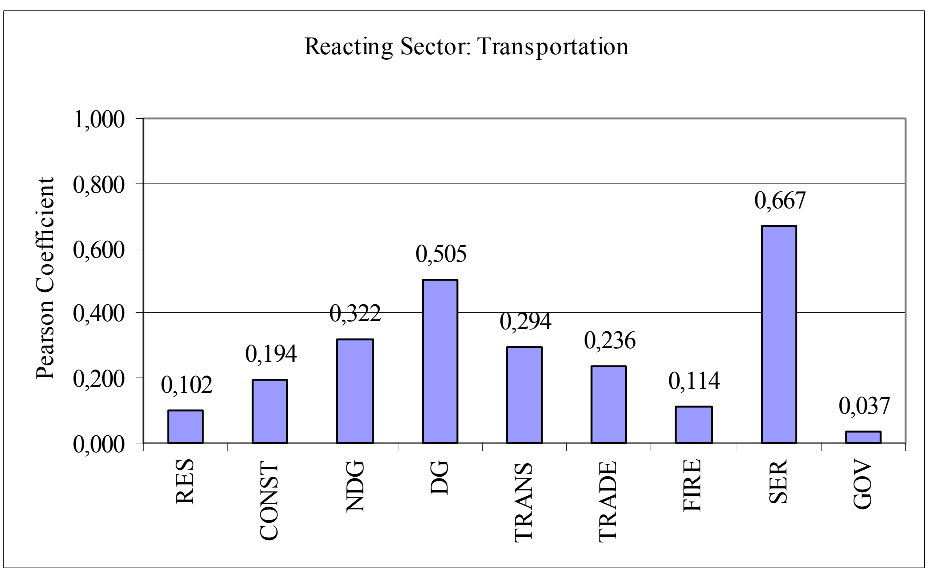

Figure 6:

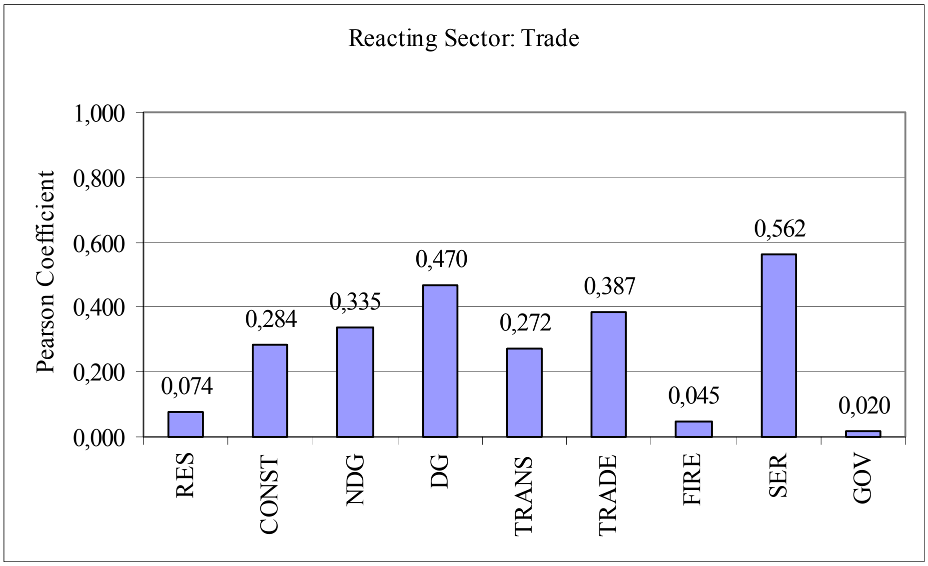

Figure 7:

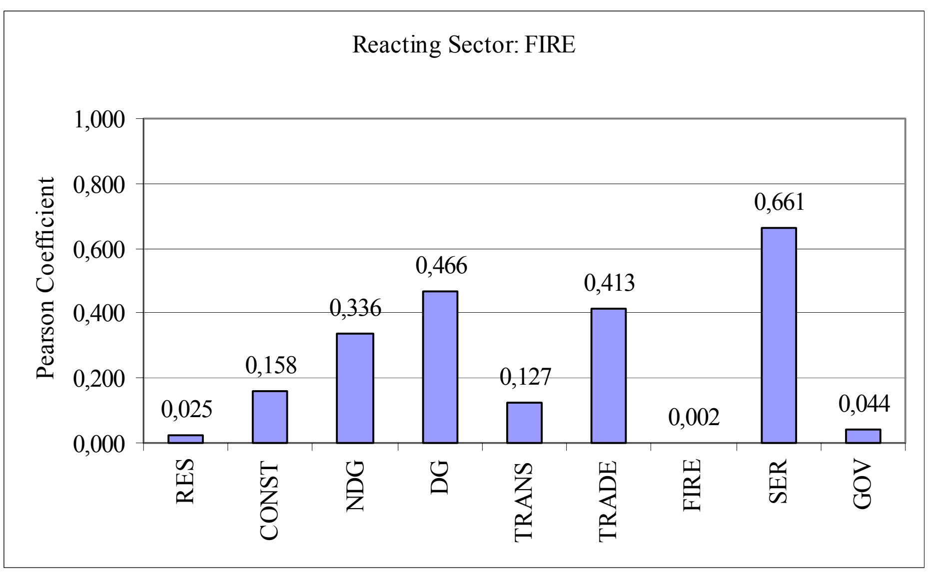

Figure 8:

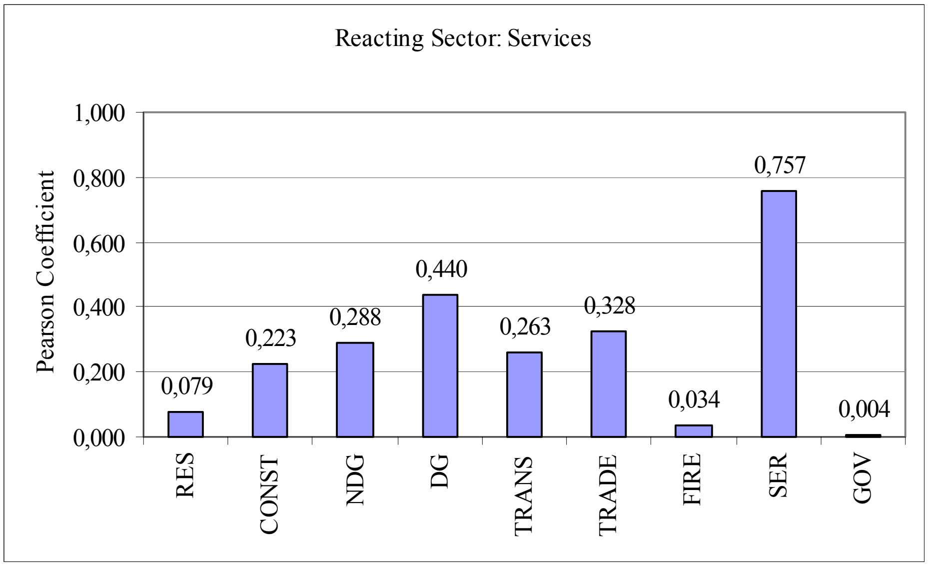

Figure 9:

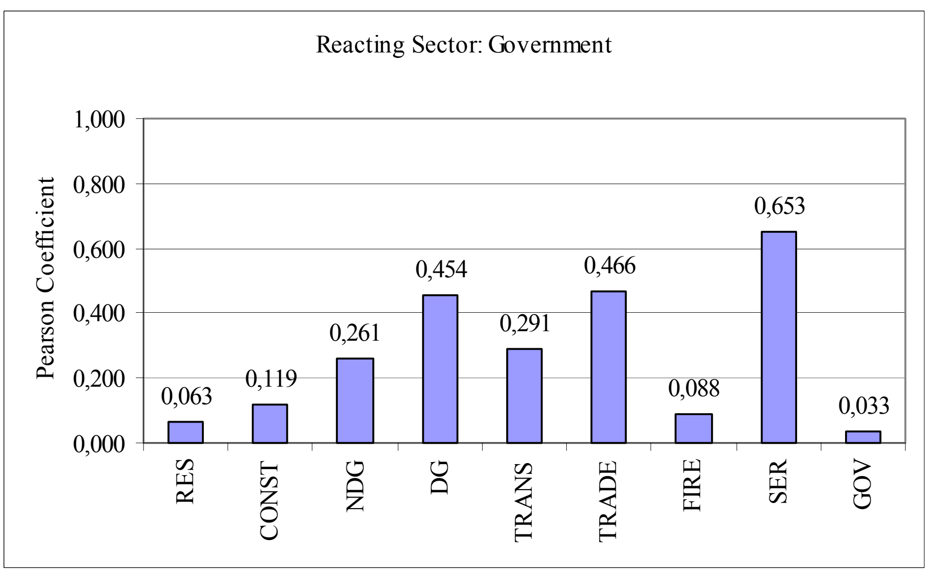

Figure 10: