SENSITIVITY OF SIMULATION RESULTS TO COMPETING SAM UPDATES

M. Alejandro Cardenete
Departamento de Economía y Empresa
Universidad Pablo de Olavide de Sevilla
41013-Sevilla SPAIN
(macarflo@dee.upo.es)

&

Ferran Sancho
Departament d'Economia
Universidad Autònoma de Barcelona
08193-Bellaterra SPAIN
(ferran.sancho@uab.es)

Barcelona Economics WP nº 88

History:
Fist draft december 2002
Revised draft october 2003

The second author wishes to express his gratitude to Catalonia's Department of Universities, Research and Information Society (DURSI) Mobility Programme BE2002-400105 for the scholarship support to visit the University of California at Berkeley where this paper has been developed and written. The research reported here is also part of the ongoing projects SEC2000-0390 from Spain's Ministry of Science and Technology and SGR2001-0164 from DURSI. All errors and shortcomings are the author's sole responsibility.

Abstract

Recently there has been a renewed research interest in the properties of non survey updates of input-output tables and social accounting matrices (SAM). Along with the venerable and well known scaling RAS method, several alternative new procedures related to entropy minimization and other metrics have been suggested, tested and used in the literature. Whether these procedures will eventually substitute or merely complement the RAS approach is still an open question without a definite answer. The performance of many of the updating procedures has been tested using some kind of proximity or closeness measure to a reference input-output table or SAM. The first goal of this paper, in contrast, is the proposal of checking the operational performance of updating mechanisms by way of comparing the simulation results that ensue from adopting alternative databases for calibration of a reference applied general equilibrium model. The second goal is to introduce a new updating procedure based on information retrieval principles. This new procedure is then compared as far as performance is concerned to two well-known updating approaches: RAS and cross-entropy. The rationale for the suggested cross validation is that the driving force for having more up to date databases is to be able to conduct more current, and hopefully more credible, policy analyses.

Keywords: Social Accounting Matrices, Input-output, Non-survey updating techniques, Applied General Equilibrium, Regional policy analysis, Evaluation of simulation results.

JEL classification: C52, C67, C68

1. Introduction

Applied General Equilibrium Analysis (AGE), as inspired by the work of Scarf (1977) and exemplified by the leading references of Dervis, de Melo & Robinson (1982), Shoven & Whalley (1984), and Ballard et al. (1986), among others, is perhaps the tool of choice when studying disaggregated resource allocation in an empirical setting. In real-world practical applications the implementation of an AGE model usually proceeds thanks to the availability of a microconsistent database known as a Social Accounting Matrix (SAM). It is theoretically possible to build and implement an AGE model without a SAM, but the operational difficulties involved in using such a procedure would render it hopeless for most practical and applied purposes. The lack of a SAM can be overcome in very small size models but even then a SAM would provide the coherent numerical background required for an efficient and effective way to proceed in the modeling effort, mainly if the model, even if small, has to match some known economic data. It is therefore strongly advisable, if not indispensable, that a SAM be available for a successful AGE model implementation. Similar considerations also apply to input-output tables and their derived models[1].

Data are used to implement empirical models and these models are then used to perform economic analyses and simulations. The quality of data and/or its currentness is therefore of critical relevance to appraise and evaluate model results and to give them credibility before policy makers and economic authorities. Unfortunately good data of the kind needed in AGE and input-output analysis is not produced in a timely and regular way by Statistical Offices. No matter how undesirable this may be from the economist's perspective, data collection and compilation is expensive both in time and resources and a temporal lag in the production and publication of official data is therefore an unavoidable reality. Ways out of this problem do exist in the form of updating techniques that permit to project forward in time a base year SAM or input-output table. This kind of SAM updating problem is a particular case of what in the linear algebra literature is referred to as a matrix balancing problem (see Rothblum & Schneider, 1989, and Schneider & Zenios, 1990). It can be stated as follows: Given an $m \times n$ non-negative base matrix $A^0 = (a_{ij}^0)$, and non-negative vectors $X^c = (X_i^c)$ in R^n and $X^r = (X_j^r)$ in R^m, find a $m \times n$ non-negative matrix $\hat{A}^1 = (\hat{a}_{ij}^1)$ close to A^0 and such that the column and row sums of the new matrix satisfy the properties $\Sigma_j \hat{a}_{ij}^1 = X_i^c$ and $\Sigma_i \hat{a}_{ji}^1 = X_j^r$.

Because of the nature of economic data, in the SAM balancing problem $m = n$ and the vectors X^c and X^r satisfy the additional restriction that $X^c = X^r = X$ (a budget restriction: total outlays, or column sums, equal total receipts, or row sums, in all sectors). The economic interpretation is that the matrices A^0 and $\hat{A}^1$ represent socioeconomic SAMs whereas the vector X describes new information on marginal totals. The most common situation is the projection of a given SAM at date t=0 to a more recent date t=1 for which a set of partial information on new marginal totals is known. A related problem is the regionalitazion of a national SAM. In this case, A^0 can be interpreted as a national or statewide SAM whereas $\hat{A}^1$ is the adjusted regional SAM for which the known marginal totals correspond to a regional decomposition of data. The conceptual structure of the problem is nonetheless the same.

The technique most commonly used in updating a SAM is the RAS or biproportional method. The origins of RAS are not clear and the technique seems to have been independently discovered several times in different fields. Schneider & Zenios (1990) report on how the RAS method has been extensively used in economics, but also in demography, probability and transportation. The appeal of the RAS procedure arises from its extremely simple algorithmic implementation. Its conceptual and mathematical properties are fully described in Bacharach (1970). More recently, entropy techniques from information theory have been adapted by Golan et al. (1994) and Robinson et al. (2001) for the updating of input-output tables and social accounting matrices. However RAS and entropy methods are closely related as Bacharach (1970, chapter 6), Schneider (1989), Schneider & Zenios (1990) and McDougall

(1999) have pointed out. RAS can also be formulated as an entropy minimization problem for total transactions but the equivalent scaling algorithm is conceptually simpler and less expensive to implement as far as programming and computing power are concerned.

As a possible alternative, or complement, to these well-known methods, we wish to introduce in this report a new approach to SAM updating that is suggested from information retrieval theory, a branch of computer science concerned with developing efficient methods of retrieving information from a data bank (Salton & McGill, 1983). Whenever a *query* for data is formulated, an algorithm fetches *documents* in a data bank that are closely related to the query in some similarity sense. The higher the similarity or matching scores between the queries and the retrieved documents, the more successful is the retrieval algorithm. A base SAM can be seen as a *query* for the ideal but unknown *document* SAM and an information retrieval algorithm will fetch from the data bank (the set of feasible SAMs) a document SAM with information content closely matching that required by the query. The nature of the algorithm is therefore based upon some concept of similarity that compares queries with documents.

The whole purpose of updating a given SAM is to solve the problem of not having an actual newer SAM. The matrix $\hat{A}^1$ is an update of the matrix A^0 but it is also an estimate or approximation to the true unknown matrix $A^1 = (a^1_{ij})$. The distance between A^0 and $\hat{A}^1$, however minimized, entails an error, unknown in magnitude if A^1 is itself unknown, between the updated matrix $\hat{A}^1$ and the true matrix A^1, as Jian (2002) has recently shown using Monte Carlo simulations. When the true matrices are finally produced and made public, it is possible to measure *ex-post* the error involved in each of the different updating procedures. This is the approach followed by Jackson & Murray (2002) who present a comprehensive statistical appraisal of the error induced by different distance minimizations.

It is this unknown error that is of concern when using an updated matrix instead of the true but unavailable one in economic modeling since errors could conceivably translate to larger than expected values in simulation results. This possible and less-than-desirable robustness phenomenon, even if it is not very likely to happen, has been theoretically pointed out by Dietzenbacher (1993) and more recently by Wolff (2002) for input-output data and models. Therefore, in any numerical model developed with the goal of performing policy analysis, the simulation results will be inevitably affected by the carried over matrix substitution. The question is not whether but how and by how much.

In dealing with empirical matrices it is quite common that the true matrices are not known and will not be known for some time. Hence, and as a proxy, the usual recourse is to perform an *ex-ante* evaluation measuring the degree of proximity between the given initial matrix and the updated ones[2].

One goal of this paper is to call the attention to the fact that checking and measuring the *ex-ante* distance performance between the base A^0 and its alternate updatings $\hat{A}^1$, and the *ex-post* error between $A^1$and $\hat{A}^1$, whenever this becomes feasible, is clearly necessary but we feel that is not sufficient. What we wish to argue in this paper is that distance and error appraisals can and should be complemented with an *ex-ante* analysis of the variability induced in simulation results by the adoption of updated data bases in place of the true unavailable one. Whenever the true matrix is made available, a similar *ex-post* appraisal could and should be of course conducted. As a first step in this direction we consider two well established updating methods, namely, RAS and cross-entropy (CE), along with the aforementioned new procedure based on information retrieval principles, to project forward in time a 1995 regional SAM of Andalusia, Spain, to known total marginals for 1999. Using the three competing SAMs, we calibrate a AGE regional tax model developed by the authors (Cardenete & Sancho, 2003) and proceed to perform a range of tax policy simulations under the three calibrated versions.

In Section 2 we succinctly present the updating techniques, the original data base and the supporting AGE regional model. Section 3 contains the numerical results and a discussion. Section 4 concludes the paper with a summary.

2. Methodology and data

2.1 The matrix balancing problem in an economic setting

The general matrix balancing problem for square matrices like SAMs which have the property that row sums coincide with column sums can be stated as follows: Let A_n be the set of $n \times n$ non-negative matrices which have no zero row or column. Let $A^0 = (a_{ij}^0) \in \mathrm{A}_n$, $X \in R^n$, and let us consider a loss function $d : \mathrm{A}_n \times \mathrm{A}_n \rightarrow R^+$. The matrix balancing problem consists in finding a matrix $\hat{A}^1 = (\hat{a}_{ij}^1) \in \mathrm{A}_n$ that solves:

$$Min\ d(A^0, \hat{A}^1)$$

subject to

(1) $\sum_{j=1}^{n} \hat{a}_{ij}^1 = X_i$ for all i

(2) $\sum_{i=1}^{n} \hat{a}_{ji}^1 = X_j$ for all j

(3) $a_{ij}^0 = 0$ implies $\hat{a}_{ij}^1 = 0$.

Restrictions (1) and (2) establish that the column and row sums of the solution matrix $\hat{A}^1$ must coincide respectively with the real values set in the common marginal vector X. Restriction (3) makes the updated matrix to inherit the zero structure of the base matrix[3]. It is the nature of the function d that gives rise to alternate updating results. Given the SAM matrix A^0 of total transactions and the new information vector X, in the RAS procedure we seek a new SAM $\hat{A}^1$ through the minimization of:

$$d(A^o, \hat{A}^1) = \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{a}_{ij}^1 \cdot (\ln \hat{a}_{ij}^1 - \ln a_{ij}^0)$$

subject to restrictions (1) to (3). The cross-entropy approach uses technical coefficient matrices in the minimand instead of total flows so that now we would be minimizing:

$$d(A^o, \hat{A}^1) = \sum_{i=1}^{n} \sum_{j=1}^{n} (\hat{a}_{ij}^1/X_j) \cdot (\ln(\hat{a}_{ij}^1/X_j) - \ln(a_{ij}^0/X_j^0))$$

subject again to (1) to (3). Here $X_j^0 = \Sigma_i\, a_{ij}^0$ is the level value for the *j*-th row and column sum in the original matrix and a_{ij}^0/X_j^0 and $\hat{a}_{ij}^1/X_j$ initial and updated technical coefficients, respectively.

In classical information retrieval theory (Salton & McGill, 1983), the performance of retrieval algorithms in vector space is evaluated using similarity indices that measure the degree of proximity, or match, between a query and a retrieved document. In general, queries and documents are represented by on/off binary properties but the similarity notion can be extended straightforwardly to continuous vectors. There is in fact a variety of similarity measures but because of its very simple mathematical structure we will exclusively focus here on the cosine similarity index. Consider any two non-negative, non-zero vectors $x, y \in R_+^n$, their inner product $< x, y >$ and the angle $\alpha(x, y)$ that these two vectors form in Euclidean space. From elementary geometry we know the following property holds:

$$< x, y >= cos\ \alpha(x, y) \cdot \| x \| \cdot \| y \|$$

where $\| x \|$ represents the Euclidean norm of vector *x* and *cos* is the cosine function. We will see now that the cosine of angle $\alpha(x, y)$ can be interpreted as a similarity between *x* and *y*. If we consider

$$cos\ \alpha(x, y) = \frac{<x, y>}{\|x\| \cdot \|y\|}$$

then the following properties hold:

a) $0 \leq \cos\alpha(x,y) \leq 1$, for all non-zero vectors $x, y \in R^n_+$

b) If $x = y$, then $\cos\alpha(x,y) = 1$ (maximum similarity)

c) If x and y are orthogonal, that is $<x, y>= 0$, then $\cos\alpha(x,y) = 0$ (maximum dissimilarity)

Property a) sets natural lower and upper bounds for the vector relationship and follows from the trigonometric definition of cosine. Property b) establishes that alike vectors have maximum similarity whereas c) says that orthogonal vectors have zero similarity. As a mater of fact property c) has a nice economic interpretation. Suppose that x and y are input requirement vectors for some output, then orthogonality means that the two technologies do not share any specific inputs ($x_i > 0$ implies that $y_i = 0$ and vice versa $y_i > 0$ implies $x_i = 0$). In this case it is all but natural that the technological similarity should be zero since the input requirement vectors are as far apart as possible in vector space and this is correctly captured by the cosine measure. The closer two vectors are, the smaller the angle they conform and the larger their similarity is.

Cosine similarity yields a proximity measure that can be used to solve the matrix balancing problem. Given a base matrix A^0 and a retrievable matrix $\hat{A}^1$ we define for each pair of columns in position j their angle $a_j(a^0_{ij}, \hat{a}^1_{ij})$. Then $\cos a_j(a^0_{ij}, \hat{a}^1_{ij})$ is a columnwise measure of the technological (or cost structure) similarity between SAM accounts j. Adding-up all column similarities to obtain a global similarity index, we can define a loss function (interpreted in this case as dissimilarity) by:

$$d(A^0, \hat{A}^1) = -\sum_{j=1}^{n} \cos a_j(a^0_{ij}, \hat{a}^1_{ij})$$

which can then be used as the minimand in the matrix balancing problem[4]. The solution of the nonlinear programming problem is thus equivalent to retrieve from A_n a feasible matrix as similar as possible to the base matrix.

2.2 The regional AGE model and database

The model was developed to represent the economy of the Spanish region of Andalusia in 1995, the most recent year for which an officially published regional input-output table was available. Using the input-output table as a backbone, the regional SAM was constructed combining information from the Regional Accounts for Andalusia developed by the regional government as well as from the general Regional Accounts elaborated by the central government operated National Statistics Institute. As it is usual in building a SAM, it was necessary to reconcile different available estimates for the same economic magnitudes. Given its wealth of micro data, the selected pivotal data source was the regional input-output table. The 1995 SAM of Andalusia consists of 37 institutional accounts, including 25 production sectors, 6 differential tax categories, 2 primary factors, a government sector, a foreign sector, a private consumption account and a capital (savings/investment) account. Unfortunately, the available 1999 data on marginals is more restrictive than the original account classification and an aggregation of the 1995 SAM has had to be performed to conform to the dimension of the new data. There are 20 accounts of which 10 are production sectors, 2 are primary factors, with 4 tax categories, as well as the customary consumption, government, foreign and capital accounts[5].

The AGE regional model is a tax policy model that follows the seminal Shoven-Whalley (1984) tradition. All markets are considered to be competitive. Technological and behavioral functions are all linearly homogeneous. Production takes place under a nesting structure. Total output is a Cobb-Douglas mix of domestic production and imports. Domestic production, in turn, combines intermediate inputs in fix proportions with a composite primary factor called value-added which in turn is a Cobb-Douglas aggregator of the two primary factors, labor and capital. Factors are assumed to be fully mobile but labor can be under-used in equilibrium giving rise to involuntary unemployment. This is accomplished in the model by way of a feedback relationship between the real wage and the unemployment level that is related to the degree of labor market flexibility.

Firms strive to maximize after-tax profits, which under the technological restriction reduces to cost minimization. All relevant conditional demand functions are obtained from the derived cost functions. On their part, consumers maximize a Cobb-Douglas utility aggregator under a disposable income budget constraint. Gross income is the result of selling endowments of labor and capital plus a set of transfers from the government and abroad. Final prices are inclusive of all indirect levies. The government collects taxes and provides social transfers, subsidies to firms and purchases public consumption and public investment. As a result of its activities, the government may incur in a running deficit if it so happens that spends more that it collects. Tax collection is governed by an endogenous tax revenue function that includes collections from an excise tax on gasoline and alcohol, a generalized value-added tax on consumption, a payroll tax levied on firms, and collections from an aggregate personal income tax.

The model is closed assuming that the level of government expenditure and export levels are given; therefore the public and foreign deficits are endogenously determined. Finally, there is an investment activity in the model whose level is determined by total private, public and foreign savings. The equilibrium concept essentially corresponds to the standard Arrow-Debreu concept for linearly homogeneous technologies, along with an additional tax revenue equilibrium condition (Ballard et al., 1985). Thus an equilibrium is a price vector, an allocation, a level of unemployment and a level of tax revenues such that all agents maximize their objective functions, all markets for goods and services clear, with the possible exception of the labor market, all taxes collected by the government equal all tax payments by all agents, and prices satisfy the unit cost rule. The existence of an equilibrium follows from the classical existence theorems. For this class of models, fortunately, uniqueness has been proved by Kehoe & Whalley (1985). Therefore meaningful comparative static exercises can be undertaken since the equilibrium set will vary smoothly with the ad-hoc, government adopted, tax structure. By modifying the tax structure and recomputing the equilibrium we can track and approximate the tax induced general equilibrium effects on the economy.

2.3 The methodological approach

Let us assume there is a policy reform to be implemented in period t=1 (or region r=1) for which it is deemed that an applied general equilibrium model is the most adequate analytical tool. Under ideal first-best conditions of full data availability, the AGE model would be calibrated using the latest available information contained in a SAM denoted by A^1. In many real-world practical applications, however, such a SAM will not exist and the only usable data base will be an older, or locationally different, A^0 SAM. In this second-best world two options are open to the modeller:

1. Calibrate the AGE model to the base A^0 SAM and interpret the simulation results *as if* they were the results of the t=1 policy taking place in t=0.
2. Calibrate the AGE model to an $\hat{A}^1$ SAM updated using one of the solutions to the matrix balancing problem and interpret the results *as if* they are close approximations to the true results (meaning those that would be derived if the true A^1 were known).

There is a tradeoff here between "timeliness" and "trueness". Under option 1 we have an old but true database and results are extrapolated to present time even if outdated. Under option 2 we have a current but only approximate database and simulation results will inevitably inherit the same characteristics. Percentage changes are clearly more credible under option 1 whereas volume or order of magnitude results are perhaps more plausible under option 2. Another consideration is that the farther apart, in time or space, are t=0 and t=1, the less satisfying are the policy simulation results of using option 1. In practice some modellers rely more on using option 1 and some on option 2 depending on the characteristics, reliability or quality of the available information.

What we propose is to examine how close are the results obtained using option 1 ("trueness" preferable over "timeliness") with those obtained using option 2 ("timeliness" preferable to "trueness") under the three different updating procedures in order to appraise to what extent SAM selection may bias the simulation results and, alternatively, which procedure may offer a more reliable, in some sense of proximity or similarity, set of simulation results. By detecting it, we may gain confidence in the interpretation of *both* percentage and volume changes induced by policy changes. In our examination, we will consider the following two tax policy scenarios:

- policy 1: a potential reform of the payroll tax (15% of effective rates),
- policy 2: a potential reform of the consumption value added-tax (30% increase of effective rates).

The justification for considering these policies rests on "harmonization" grounds. In the first case, payroll tax rates in Spain are among the highest in the European Union, hence a reduction would lead the economy towards "European" rates. In the second case, an increase in the consumption value-added tax would make Spanish tax rates closer to average "European" rates. In addition, both tax policies would be expected to have broad allocative effects, and thus a general equilibrium analysis is an appropriate research instrument.

3. Discussion

In this section we present two types of comparisons. First, we proceed to compare the SAMs that are obtained by solving the matrix balancing problem in terms of proximity measures. Second, we direct our attention to the simulations results that would ensue from adopting each candidate SAM as the numerical backbone of an applied general equilibrium model. The first set of comparisons can be thought of as model input comparisons since the derived SAMs are the necessary input for implementing the AGE model, whereas the second set of data can be referred to as model output comparisons since they use a set of simulation results produced by the model.

3.1 Comparing inputs to the model

The balancing problem takes the regional 1995 SAM as the base A^0 matrix and uses a vector of marginals X for 1999 that have been obtained from the official regional product and income accounts. Three alternative $\hat{A}^1$ SAMs are produced using RAS, cross-entropy and cosine similarity as minimands of the balancing problem. We shall refer to them, respectively, by SAMRAS, SAMCE, and SAMCOS. Table 1 present some summary proximity (distance) indicators for both coefficients and transactions between the original 1995 SAM and the three 1999 projections which, to avoid any implicit numerical bias, are unrelated to the used loss functions. The included proximity statistics include the well-known standard percentage error (STPE), Theil's U, and Lahr's (2001) weighted absolute difference (WAD):

$$STPE(A^0, \hat{A}^1) = \frac{\Sigma_i \Sigma_j \; |a_{ij}^0 - \hat{a}_{ij}^1|}{\Sigma_i \Sigma_j \, a_{ij}^0}$$

$$U(A^0, \hat{A}^1) = \sqrt{\frac{\Sigma_i \Sigma_j \left(a_{ij}^0 - \hat{a}_{ij}^1\right)^2}{\Sigma_i \Sigma_j \, a_{ij}^2}}$$

$$WAD(A^0, \hat{A}^1) = \frac{\Sigma_i \Sigma_j \, a_{ij} \cdot | a_{ij}^0 - \hat{a}_{ij}^1 |}{\Sigma_i \Sigma_j \, a_{ij}^0}$$

For the three measures, smaller numbers indicate better performance but they do not have, however, any natural upper bounds. Therefore, to facilitate comparisons we will report them in terms of relative performance arbitrarily fixing the lowest one equal to unity.

We will also consider an index of similarity introduced by Le Masné (1990) to compare column coefficients in input-output tables and that we extend here to the overall SAM coefficient structure. Le Masné's index lies between 0 and 1, with the closer the index being to 1 the larger being the similarity. We use a simple arithmetic mean of the *n* SAM columns proximities:

$$LM(A^0, \hat{A}^1) = \frac{1}{n} \cdot \sum_{j=1}^{n} (1 - 0.5 \cdot \sum_{i=1}^{n} \; | \; a_{ij}^0 - \hat{a}_{ij}^1 \; |)$$

Finally, we compute the standard Pearson's correlation coefficient as a measure of linear closeness between SAMs. Unlike the previous measures, both Le Masné and Pearson's indices have well defined bounds and are reported at their actual numerical values.

In the coefficients sub-block of Table 1 we observe than RAS performs worse than both cross-entropy and cosine similarity under all measures. Cross-entropy and cosine similarity indicators are quite close to each other but the first one dominates in four out the five indicators. When we look at the transactions sub-block, however, the situation reverses. RAS performs best in all cases; cross-entropy and cosine similarity are again somewhat close to each other and each measure dominates in two out of the four cases. These results seem to give support to the assertion by Robinson et al (2001), in their vis a vis comparison between RAS and cross-entropy methods, that RAS is generally best for projecting transactions matrices whereas cross-entropy does a better job when the interest lies in projecting coefficient matrices. On the other hand, the indices suggest that cosine similarity gives rise to a more middle-of-the-road, compromise solution.

Table 1: Proximity measures to SAM95

	SAMRAS	SAMCE	SAMCOS
Coefficients			
STPE	1,3571	**1**	1,0931
U (Theil)	1,2274	**1**	1,0242
WAD	1,3516	**1**	1,1955
Le Masné	0,8563	**0,8941**	0,8843
Pearson	0,9606	0,9732	**0,9737**
Transactions			
STPE	**1**	1,0336	1,0869
U (Theil)	**1**	1,0936	1,0900
WAD	**1**	1,0448	1,1664
Pearson	**0,9297**	0,9145	0,9211

The good news from the data in Table 1 is that there is a clear categorization among the three alternate SAMs in terms of their coefficient or flow proximities to the base SAM A^0. The not so good news, however, is that there is no conclusive recommendation as to what SAM should be selected in implementing an AGE model. On the one hand, the matrix of input-output coefficients plays a relevant role in the commodity equilibrium conditions, which may suggest a preference for using SAMCE; on the other hand, main results are commonly reported in aggregate transaction terms and this perhaps hints at SAMRAS as preferable. Or perhaps, relying on convexity, as economists would say, we could decide on using SAMCOS since there is really no way of knowing the a priori appropriate mix between coefficient and transaction preferability in any given model implementation. Therefore an additional check for further evidence is called for.

3.2 Comparing outputs from the model

Each projected SAM is a candidate database to be used in the calibration of the general equilibrium model. A set of calibrated production technical coefficients, utility coefficients and effective tax rates, specific to each of the SAMs, can be produced in such a way that once substituted into the behavioral and structural model equations they generate a benchmark equilibrium. For each equilibrium, we calculate its GDP and its associated decomposition into its standard income and expenditure sides both in absolute and percentage level. We also present the income and expenditure government accounts. Table 2 summarizes the results. Unlike the base equilibrium for 1995, which is expressed in current prices, all projected benchmark values are expressed in terms of the selected numeraire (the wage rate).

The benchmark results are seen to be strikingly robust to SAM selection. Even considering that AGE models are not prediction models and that aggregation works to smooth out differences, the degree of similarity among the results generated by adopting the three alternate databases is noteworthy and reassuring as far as confidence on performing policy simulations is concerned. This is particularly remarkable when we look at the composition of government tax revenues, a crucial check for the credibility of the simulation results that will be derived from a tax policy model.

Table 2: Aggregate indicators for base SAM 95 and projected SAM 99 Benchmarks

	SAM95	%GDP	SAMRAS	%GDP	SAMCE	%GDP	SAMCOS	%GDP
MACRO INDICATORS								
Wages and salaries	3.190.651	0,3480	4.043.008	0,3354	4.043.008	0,3356	4.043.008	0,3364
Business income	4.684.521	0,5109	5.965.350	0,4948	5.965.349	0,4951	5.965.350	0,4964
Net indirect taxes	1.293.851	0,1411	2.047.475	0,1698	2.039.981	0,1693	2.009.248	0,1672
GDP-Income	**9.169.023**	**1,0000**	**12.055.833**	**1,0000**	**12.048.338**	**1,0000**	**12.017.606**	**1,0000**
Private consumption	6.276.539	0,6845	7.892.806	0,6547	7.938.697	0,6589	7.639.948	0,6357
Investment	2.554.606	0,2786	4.094.765	0,3397	4.094.767	0,3399	4.094.765	0,3407
Public consumption	2.001.000	0,2182	2.765.039	0,2294	2.731.769	0,2267	2.730.950	0,2272
Trade balance	-1.663.122	-0,1814	-2.696.777	-0,2237	-2.716.894	-0,2255	-2.448.057	-0,2037
GDP-Expenditure	**9.169.023**	**0,9999**	**12.055.833**	**1,0001**	**12.048.339**	**1,0000**	**12.017.606**	**0,9999**
GOVERNMENT								
Net Production Taxes	-422.658	-0,0461	-383.093	-0,0318	-390.587	-0,0324	-421.320	-0,0351
VAT	597.476	0,0652	897.807	0,0745	897.808	0,0745	897.807	0,0747
Payroll tax	1.119.033	0,1220	1.532.761	0,1271	1.532.761	0,1272	1.532.761	0,1275
Income Tax	933.719	0,1018	1.232.508	0,1022	1.232.508	0,1023	1.232.508	0,1026
Public spending	4.092.415	0,4463	5.018.894	0,4163	5.011.399	0,4159	4.980.667	0,4144
Public Deficit	-1.864.845	-0,2034	-1.738.909	-0,1442	-1.738.908	-0,1443	-1.738.909	-0,1447

In Tables 3A to 3D we describe the outcome of simulating a 15 percent decrease in the payroll tax using the AGE model implemented with the old but true SAM for 1995 and then with the three new but estimated 1999 SAMs. Similarly, Tables 4A to 4D summarize the results of a 30 percent increase in VAT rates. Relative prices, activity levels, aggregate macro values and government and welfare indicators are displayed. The selection of units in each benchmark equilibrium yields prices and activity levels initially normalized to unity, thus any deviation from the unitary initial value also indicates the percentage change in the variable. For relative prices it is understood that the change is in terms of the price of the numeraire good. Changes in sectoral activity levels and the unemployment rate, however, can easily be interpreted as physical changes. All aggregate, government and welfare variables are ratio indicators and thus do not depend on the chosen numeraire.

The basic message that emerges from the analysis of the policy simulations results is once again one of overall robustness. In Table 3A, for instance, the biggest price impact is detected consistently in sector 8 whereas the smallest effect takes place in sector 9 in all four cases. It can also be seen that if sectors were ordered from largest to smallest price impact, the resulting sectoral ordering would be essentially the same, except for one sector switch (sectors 3 and 6 under SAMCOS), under all three 1999 estimated databases. As for activity levels, Table 3B shows that under all benchmarks the recipient sectors with the highest and lowest impact are the same (sectors 5 and 10, respectively). Ordering sectors according to impact, however, is not as robust as it turns out to be with relative prices. The average impact on prices and activities, as measured by a Consumers Price Index and an Industrial Activity index, are for all practical purposes the same under all databases.

Table 3A. Relative Prices after a 15% decrease in Payroll Tax

SECTORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
1. Agriculture, cattle, forestry and fishing	0,9823	0,9821	0,9819	0,9813
2. Extractives	0,9827	0,9821	0,9821	0,9814
3. Energy	0,9852	0,9841	0,9846	0,9836
4. Manufactures	0,9830	0,9823	0,9822	0,9809
5. Construction	0,9785	0,9752	0,9758	0,9735
6. Commerce	0,9837	0,9837	0,9839	0,9839
7. Transportation & communications	0,9829	0,9824	0,9824	0,9817
8. Other services	0,9746	0,9728	0,9728	0,9718
9. Commercial services	0,9931	0,9933	0,9932	0,9951
10. Non commercial services	0,9771	0,9753	0,9760	0,9750
Consumer Price Index	0,9837	0,9832	0,9829	0,9835

Table 3B. Activity levels after a 15% decrease in Payroll Tax

SECTORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
1. Agriculture, cattle, forestry and fishing	1,0117	1,0074	1,0074	1,0095
2. Extractives	1,0151	1,0107	1,0132	1,0127
3. Energy	1,0175	1,0200	1,0203	1,0230
4. Manufactures	1,0152	1,0134	1,0119	1,0127
5. Construction	1,0255	1,0225	1,0206	1,0250
6. Commerce	1,0214	1,0217	1,0211	1,0218
7. Transportation & communications	1,0161	1,0142	1,0147	1,0136
8. Other services	1,0132	1,0180	1,0200	1,0177
9. Commercial services	1,0144	1,0145	1,0146	1,0116
10. Non commercial services	1,0009	1,0016	1,0029	1,0030
Industrial Activity Indicator	1,0160	1,0164	1,0166	1,0163

Table 3C: Macroeconomic indicators after a 15% decrease in Payroll Tax

INDICATORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
Unemployment Rate change	-0,0220	-0,0252	-0,0257	-0,0247
Wages and Salaries/GDP	0,3606	0,3483	0,3486	0,3492
Business Income/GDP	0,5126	0,4965	0,4967	0,4982
Net Indirect Taxes/GDP	0,1268	0,1552	0,1547	0,1526
Private Consumption/GDP	0,6915	0,6624	0,6665	0,6431
Investment/GDP	0,2810	0,3408	0,3404	0,3420
Public Consumption/GDP	0,2137	0,2244	0,2218	0,2224
Trade Balance/GDP	-0,1863	-0,2275	-0,2287	-0,2075

Table 3D: Government and Welfare indicators after a 15% decrease in Payroll Tax

INDICATORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
Net Production Taxes/GDP	-0,0463	-0,0319	-0,0324	-0,0351
VAT/GDP	0,0653	0,0746	0,0746	0,0749
Payroll Taxes/GDP	0,1077	0,1124	0,1125	0,1128
Income Taxes/GDP	0,1029	0,1034	0,1035	0,1038
Public Deficit/GDP	-0,2091	-0,1504	-0,1503	-0,1509
Welfare change/Tax Revenues	0,1088	0,0961	0,0973	0,0978
Welfare change/GDP	0,0250	0,0248	0,0251	0,0251
Marginal Welfare change	-1,7975	-1,7203	-1,7547	-1,7411

Table 4A: Relative Prices after a 30% increase in VAT

SECTORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
1. Agriculture, cattle, forestry and fishing	1,0039	1,0082	1,0049	1,0014
2. Extractives	1,0157	1,0178	1,0158	1,0093
3. Energy	1,0188	1,0242	1,0200	1,0181
4. Manufactures	1,0193	1,0219	1,0205	1,0106
5. Construction	1,0235	1,0351	1,0363	1,0323
6. Commerce	1,0005	0,9985	0,9991	1,0006
7. Transportation & communications	1,0205	1,0266	1,0238	1,0208
8. Other services	1,0073	1,0099	1,0118	1,0160
9. Commercial services	1,0016	1,0043	1,0040	1,0034
10. Non commercial services	0,9980	0,9957	0,9972	0,9958
Consumer Price Index	1,0079	1,0082	1,0076	1,0074

Table 4B: Activity Levels after a 30% increase in VAT

SECTORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
1. Agriculture, cattle, forestry and fishing	0,9949	0,9960	0,9968	1,0001
2. Extractives	0,9937	0,9971	0,9938	0,9897
3. Energy	0,9881	0,9797	0,9787	0,9720
4. Manufactures	0,9895	0,9878	0,9930	0,9973
5. Construction	0,9922	0,9880	0,9958	1,0012
6. Commerce	0,9886	0,9890	0,9899	0,9889
7. Transportation & communications	0,9879	0,9842	0,9846	0,9877
8. Other services	0,9930	0,9893	0,9877	0,9899
9. Commercial services	0,9879	0,9835	0,9842	0,9846
10. Non commercial services	0,9997	0,9995	0,9991	0,9991
Industrial Activity Indicator	0,9909	0,9891	0,9904	0,9916

Table 4C: Macroeconomic indicators after a 30% increase in VAT

INDICATORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
Unemployment Rate	0,0104	0,0119	0,0111	0,0107
Wages and Salaries/GDP	0,3421	0,3289	0,3291	0,3302
Business Income/GDP	0,4997	0,4825	0,4828	0,4837
Net Indirect Taxes/GDP	0,1582	0,1886	0,1881	0,1860
Private Consumption/GDP	0,6749	0,6435	0,6474	0,6246
Investment/GDP	0,2820	0,3442	0,3468	0,3488
Public Consumption/GDP	0,2188	0,2293	0,2268	0,2276
Trade Balance/GDP	-0,1757	-0,2170	-0,2210	-0,2010

Table 4D: Government and Welfare indicators after a 30% increase in VAT

INDICATORS	SAM95	SAMRAS99	SAMCE99	SAMCOS99
Net Production Taxes/GDP	-0,0459	-0,0314	-0,0322	-0,0349
VAT/GDP	0,0841	0,0953	0,0956	0,0957
Payroll/GDP	0,1199	0,1247	0,1247	0,1252
Income Taxes/GDP	0,1004	0,1005	0,1005	0,1008
Public Deficit/GDP	-0,1899	-0,1282	-0,1281	-0,1287
Welfare change/Tax Revenues	-0,0934	-0,0885	-0,0848	-0,0817
Welfare change/GDP	-0,0242	-0,0256	-0,0245	-0,0234
Marginal Welfare change	-1,5058	-1,4282	-1,3371	-1,2827

The direct and indirect effects in labor costs that would follow from the adoption of this tax policy reduction would translate to a fall in the unemployment rate of about 2.5 percent points using any of the estimated SAMs, not substantially higher that the 2.2 percent reduction that the model with the 1995 database foresees. Similar conclusions can be drawn from the resemblance among the rest of aggregate variables as seen in Table 3C which, with the noted exception of private consumption and the trade deficit under SAMCOS, are noticeably similar.

The very important welfare variables show again a high degree of robustness. We can see in Table 3D that there would be a welfare improvement following the payroll tax reduction policy of about 9.7 points over total tax revenue and of about 2.5 points over GDP. The reported marginal welfare change has been computed as a numerical derivative and measures the ratio of welfare increase that would result from the tax decrease, hence its negative sign. All three derivatives have a very similar value which lies slightly under the 1995 estimate, but not by any large quantity.

Tables 4A to 4D report the same variables for the second policy scenario where a 30 percent tax increase in VAT rates would be enacted. Under this tax policy prices tend to increase relative to the numeraire and activity levels tend to fall. There is no substantial difference among the estimated average impact on prices and activity levels. Highest and lowest price impacts are shared by all versions of the model (sectors 5 and 10, respectively). In activity levels, the lowest impact is always detected in sector 3. The sector with the most impact identified by the SAMCOS model (sector 5) departs, however, from the sector with the highest impact pinpointed by the other model versions (sector 10). Aggregate variables as well as government and welfare indicators show a close degree of resemblance. Interestingly, the marginal welfare change is smaller in absolute value with VAT than with the payroll tax, an indication that there possibly is some margin for efficiency gains if an appropriately designed, revenue neutral tax reform would be enacted.

The data displayed in the previous Tables tentatively suggest that there is a high level of robustness in the simulation outcomes. Despite the intrinsically different nature of the updating procedures, all seem to yield SAM matrices that give rise to general equilibrium outcomes very close to each other. In a sense this is quite reassuring but it is not by any means a proof of reliability since it is basically a subjective impression of what certainly seems to be a good match. The key variables in a general equilibrium model are prices and quantities. Once we know them all the other variables and indicators can be calculated quickly and easily. Therefore we could go the root equilibrium data and compare prices and activity levels to obtain a measure of goodness of fit between the base equilibrium data for 1995 and those derived from the three projected SAMs.

Economic theory tells us that in terms of prices all a general equilibrium model can produce are relative prices. Even though it is customary to report changes in the CPI, a word of caution is needed since any obtained value will always depend on the chosen numeraire. General equilibrium models cannot say anything about inflation or absolute price changes. Consequently, to control for this situation we will deflate all prices so that they yield a unitary CPI index. This is achieved by dividing all prices by the pre-adjustment CPI. With this procedure we eliminate spurious price growth by restricting commodity prices to belong to a weighted unit simplex.

We will use two measures of goodness of fit to appraise how well simulation prices and activity levels from the three 1999 model version match prices and activity levels from the base 1995 model. The first descriptive statistic is the well-known Pearson correlation coefficient, an index that captures the relative direction and relative magnitudes of the 1999 "predicted" effects vis a vis the 1995 results. The second statistic will be a weighted correlation coefficient that captures deviations between the predictions obtained calibrating the model with the base SAM A^0 and each of the three alternative SAMs $\hat{A}^1$:

$$\rho(A^0,\hat{A}^1) = \left(\sum_{i=1}^{n} \beta_i^2 \cdot y_i^0 \cdot \hat{y}_i^1\right) / \sqrt{\left(\sum_{i=1}^{n} \beta_i^2 \cdot (y_i^0)^2 \cdot \sum_{i=1}^{n} \beta_i^2 \cdot (\hat{y}_i^1)^2\right)}$$

In this expression β_i is the relative size of sector *i*, y_i^0 is the change in the variable in sector *i* when compared to pre-simulation benchmark values, and $\hat{y}_i^1$ is the change in variables predicted by the alternative model versions. In the relative price comparison the weights correspond to the CPI weights, whereas for activity levels we have used value-added shares. The results are presented in Table 5.

Table 5: Relative prices and activity levels indicators

	SAMRAS	SAMCE	SAMCOS
Prices: 15% payroll tax decrease			
Pearson	0,9915	**0,9943**	0,9858
Weighted correlation	**0,9931**	0,9869	0,9683
Prices: 30% VAT increase			
Pearson	**0,9799**	0,9612	0,8632
Weighted correlation	**0,9910**	0,9895	0,8784
Activities: 15% payroll tax decrease			
Pearson	**0,8971**	0,8385	0,8919
Weighted correlation	0,9896	0,9786	**0,9898**
Activities: 30 % VAT increase			
Pearson	**0,8889**	0,7885	0,6195
Weighted correlation	**0,9825**	0,9484	0,9055

An examination of the table shows that all correlations are quite high and that the model implemented with the SAM updated with RAS yields in general the best match to 1995 results. Cross-entropy and cosine similarity dominate only in one category each. Cross-entropy comes in second position when we look at prices, whereas when we look at activity levels cross-entropy and cosine similarity alternate in second position depending on the tax simulation. As for categories, relatives prices seem to have a slightly higher prediction power than that of quantities. Even though these correlations coefficients are all closer to 1, therefore in the good side of the interval, it should be remembered that they do not indicate any causal links. Their high value should perhaps be considered as a necessary, but not sufficient, evidence of good modeling practices. Their high value probably reflects that the underlying model structure is common to all versions and that the alternative databases are constructed using also a common initial matrix and a common vector of updated marginals. What in fact matters most for a performance evaluation is their relative size.

4. Concluding Remarks

We have reported in this paper a limited, empirical test of the performance of three different means of solving the matrix balancing problem in an economic setting. Instead of focusing exclusively on a comparison of the resulting updated matrices, we have also examined the implications of adopting each competing SAM in the calibration and implementation of an applied general equilibrium model that has been used to simulate two wide-ranging tax policy reforms in a regional economy. The overall impression is that economic results are not very sensitive to the choice of updated database, and not very different in turn to results in the base year. This is clearly a reassuring conclusion since it indicates that we are not consistently off-mark when carrying out policy evaluation analysis. This observation can also be taken in two ways. Firstly, if the time difference between the true available SAM and the period of interest is not too large, perhaps confining the analysis to the old database is not too restrictive. An added benefit is that usually survey SAMs have a higher level of disagreggation than non-survey ones, thus providing a finer degree of detail in the microeconomic results. But, secondly, if macro aggregates command a higher value to policy makers, then using an updated SAM could provide a better, more up-to-date answer.

Another conclusion is that an *ex-ante* examination of distance measures among alternative SAMs is not necessarily an indication of the *ex-post* performance of simulations.

Recall that RAS gave better distance indicators when looking at transactions while cross-entropy came ahead when looking at coefficient matrices, with cosine similarity being second on both counts. However, the simulation results hint at a possible RAS dominance over the other two methods when evaluated using correlation coefficients.

More thorough testing is obviously needed. Our experiment is a one-shot experiment using a specific database and a specific vector of marginals and no extrapolations to general conclusions should be drawn at this stage. Further research is clearly required and a possibility for a more systematic testing is using Monte Carlo simulations where instead of actual marginal vectors, randomly selected ones could be used.

More testing and perhaps refinements of the information retrieval similarity approach is also necessary. Although orthogonality has an interesting economic interpretation, the cosine function does not seem to outperform RAS or cross-entropy, coming in second to RAS when transactions matrices matter most and second to cross-entropy when coefficients matter most. This middle-of-the road performance, however, could be useful when and if there is no clear cut preference over transactions or coefficients. To learn more about the properties of similarities, the use of the cosine function should be complemented with some of the other similarities indicators and their *ex-ante* (distance) and *ex-post* (simulation) performance compared among themselves and again with the standard RAS and cross-entropy methods.

5. References

Bacharach, M. (1970), *Biproportional Matrices and Input-output Change*, Cambridge University Press, London.

Ballard, C., D. Fullerton, J. Shoven and J. Whalley (1985), *A General Equilibrium Model for Tax Policy Evaluation*, NBER and University of Chicago Press, Chicago.

Cardenete, M. A. and F. Sancho (2003), "A General Equilibrium Model for Tax Policy Evaluation", *Review of Urban and Regional Studies*

Dervis, K., J. de Melo and S. Robinson (1982), *General Equilibrium Models for Development Policy*, Cambridge University Press, New York.

Dietzenbacher, E. (1990), "The Sensitivity of Input-output Multipliers", *Journal of Regional Science*, 30(2), pp. 239-58.

Jian, X. (2002), "Distance, Degree of freedom and Error of RAS Method", working paper presented at the 14th International Conference on Input-output Techniques, Montreal.

Jackson, R. and T. Murray (2002), "Alternate Methods for Minimizing Inter-Matrix Distances", Research paper 2002-9, Regional Research Institute, West Virginia University.

Golan, A., G. Judge and S. Robinson (1994), "Recovering Information from Incomplete or Partial Multisectoral Economic Data", *Review of Economics and Statistics*, 76, pp. 541-9.

Kehoe, T. and J. Whalley (1985), "Uniqueness of Equilibrium in Large Scale Numerical General Equilibrium Models", *Journal of Public Economics*, 28, pp. 247-54.

Lahr, M. (2001), "A Strategy for Producing Hybrid Regional Input-output Tables", in *Input-Output Analysis: Frontiers and Extensions*, edited by E. Dietzenbacher and M. Lahr, Palgrave, pp. 211-42, New York.

Le Masné, P. (1990), "Le Systeme Productif Français face a ses Voisins Europeens", in *La Comptabilité nationale face au defí international*, edited by E. Archanbault and O. Arkhipof, Economica, Paris.

McDougall, R. (1999), "Entropy Theory and RAS are Friends", GTAP working paper No. 6, Purdue University.

Robinson, S., A. Cattaneo and M. El-Said (2001), “Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods”, *Economic Systems Research*, 13(1), pp. 47-64.

Rothblum, U. and H. Schneider (1989), “Scaling of Matrices Having Preespecified Row Sums and Column Sums”, *Linear Algebra and its Applications*, 114/115, pp. 737-64.

Salton, G, and M.J. McGill (1983), *Introduction to Modern Information Retrieval*, McGraw-Hill, New York.

Schneider, H. (1989), “Matrix Scaling, Entropy Minimization and Conjugate Duality (I): Positivity Conditions”, *Linear Algebra and its Applications*, 114/115, pp. 785-813.

Shoven, J. and J. Whalley (1984), “Applied General Equilibrium Models of Taxation and International Trade”, *Journal of Economic Literature*, 22(3), pp. 1007-51

Schneider, M. and S. Zenios (1990), “A Comparative Study of Algorithms for Matrix Balancing”, *Operations Research*, 38(3), pp. 439-55

Thissen, M. and H. Logfren (1998), “A New Approach to SAM updating with an application to Egypt”, *Environment and Planning A*, 30, pp. 1991-2003.

Wolff, R. (2002), “A Global Robustness Measure for Input-output Projections from ESA and NSA tables”, working paper presented at the 14th International Conference on Input-output Techniques, Montreal.

Endnotes

[1] An input-output table can always be embedded in a SAM as a data subset. When no need for clarification is required we will refer to SAM updating as a term encompassing both the updating of a SAM proper and an input-output table.

[2] This is the approach followed by Thissen and Logfren (1999) and Robinson et al (2001).

[3] Consider the unrestricted possibility that a zero entry would become positive after the updating. There are two types of zero entries in a SAM. The first type are "technical" zeros; that is the case of an input that is not actually used in the production of a commodity but could be used under a different technology. The second type of zeros are "conceptual" zeros. In no SAM labor is directly used in the production of capital, nor in generating excise taxes. Hence the updating cannot be allowed to change these second type zero entries into positive or negative values. We believe the safest course of action is to maintain the initial matrix zero structure.

[4] Another widely used measure of similarity is Jaccard similarity (Salton and McGill, 1983). The Jaccard index in vector space is defined as

$$\mathrm{Jac}(x,y) = \frac{<x,y>}{\sum x_i^2 + \sum y_i^2 - <x,y>}$$

It can be seen that it satisfies the same properties as cosine similarity. This measure is commonly used in the biological sciences to compare populations.

[5] Both SAMs are available from the authors upon request.