A Generalized Cross Entropy formulation for matrix balancing with both positive and negative entries

Esteban Fernandez-Vazquez[1]

RegioLab and Department of Applied Economics

University of Oviedo, Spain

Abstract

This paper presents a matrix balancing technique based on Generalized Cross Entropy (GCE) that can be suitable for matrices containing both positive and negative entries. This technique makes possible sing flips in the cells of the initial and the estimated matrices, which can be something desirable in situations where assuming sign-prevention for all the entries of the matrix could be too restrictive. An additional advantage is that GCE allows for doing some inference with the estimates, something not possible when using biproportional balancing techniques like Generalized RAS (GRAS), which is the method commonly applied to balance matrices with positive and negative cells. The basic idea of the proposed GCE method is to assume each cell of the target matrix as a random variable for which we have partial information in the initial matrix. The GCE procedure assumes each observation in this matrix as a specific realization of a random process that generates the cells and it requires setting exogenously some bounds for the maximum and minimum values that this random process could generate. From this information, together with some partial data on the target matrix, the adjustment process is approached as a -constrained- minimization problem of a Kullback-Leibler divergence. A simple illustrative example shows how GCE works when adjusting a matrix characterized by having positive and negative entries within a Supply and Use (SUT) framework. Additionally, its performance is evaluated by means of a numerical simulation.

[1] The author gratefully acknowledges the financial support by the grant ECO2013-48161-R from the Spanish Ministry of Economy and Competitiveness. The usual disclaimer applies.

1. Introduction

The information contained in an Input-output (IO) table summarizes the interactions that characterize a certain economy. IO tables and Social Accounting Matrices (SAM's) are an essential quantitative tool both economic researchers and policy makers, given the huge number of possible applications, including Computable General Equilibrium (CGE) models, for analyzing productivity across industries, evaluating tax policies, studying income inequality, calculating global value chains in international trade or investigating environmental issues (see Miller and Blair, 2009, for a recent and exhaustive list of potential applications). IO datasets based on detailed surveys are, however, expensive and time-consuming for the statistical agencies. As a consequence, most of countries produce survey-based IO tables characterized by not being available for every year but every several years – intervals of five years is the most common situation-. Additionally, there normally is a considerable lag between the moment of the data collection and the publication of the IO table. In this context, the use of some non-survey method for estimating IO matrices becomes useful if not necessary in many cases. Many different non-survey methods can be employed, but all they basically consist in some type of matrix adjusting or balancing: the general solution is a matrix that diverges least with respect to some prior matrix while being consistent with some aggregate observable information.

The well-known biproportional RAS adjustment lies within this general problem and is the most frequently applied technique if all the cells in the matrix are positive. Adjusting a matrix with both positive and negative entries, however, implies some practical problems for RAS. If a RAS adjustment is applied to a matrix that contains negative cells, it easily leads to a solution that may largely deviate from the structure of the prior matrix. Junius and Oosterhaven (2003) proposed the so-called generalized RAS (GRAS) as an alternative adjustment for such situations.[2] The original GRAS

[2] As Temurshoev et al. (2013) point out, the proposal by Junius and Oosterhaven (2003) actually bases on a previous methodology developed by Günlük-Senesen and Bates (1988).

formulation is a sign-preserving technique for adjusting a matrix, and it can be applied directly to positive and negative cells. This technique, as well as several other sign-preserving adjusting techniques, defines an objective function written in terms of absolute values with respect to the matrix entries.

GRAS and its variants are the most popular techniques for matrix balancing given its computational simplicity, but their sign-preserving nature can be problematic in some cases. Consider, for example, the case of an initial matrix with all its entries in a particular row being positive (negative) whose elements are going to be adjusted to make them consistent with some posterior aggregate which is negative (positive). This is a typical case labelled as GRAS-infeasible, since the original GRAS formulation cannot deal with such a situation. Recent papers have addressed this issue, like Temurshoev et al. (2013) and Lenzen et al. (2014), which in a similar fashion proposed modifications in the GRAS algorithm to allow for sing flips in the elements to be adjusted. These techniques are recommended for dealing with these infeasibilities and to find a balanced solution by means of GRAS.

But non-preserving the sign of an element in the initial matrix could be desirable even when the balancing problem remains feasible. This could happen when one or several elements of the initial matrix could have actually sign flipped, but the aggregate data that act as a constraint is compatible with the initial sign.[3] In such a situation the mentioned techniques that modify GRAS to make it non sign-preserving will not be applied because the balancing problem is GRAS-feasible. However, accounting for the changes in

[3] For an example, consider the case of an initial input-output matrix like the one depicted in Table 1 (taken from Junius and Oosterhaven, 2003, page 94), where the total output in one industry is given by the intermediate and final demand plus net exports. The cell 'next exports' in the first row of Table 1 is negative (-3) and a GRAS adjustment will preserve this negative sign, even when in the -unknown- target matrix this cell could have turned into positive.

the sign of this set of elements would be desirable since it would improve the accuracy of the estimated target matrix.

In this paper a Generalized Cross-Entropy (GCE) estimation is proposed for this type of problems. One of the advantages of the proposed technique is that it introduces more flexibility in the adjustment and that it allows for potential changes in the sign of the cells, if the researcher considers that not preserving the sign in some of the cells of the matrix to be balanced could make sense - even when the adjustment problem remains feasible to be solved by means of conventional GRAS-. In other words, the technique suggested here assumes that a change in the sign of the entries could be improbable but not absolutely impossible. Additionally, applying GCE allows for doing some inference with the estimates, which is not possible when using some biproportional balancing technique.

The paper is divided into four additional sections. Section two presents the basic formulation of the proposed technique, whereas section three shows its solution and several inference and diagnosis tools that can be implemented by using GCE in this context, together with an illustrative example. Section four compares the performance of the GCE estimation with other adjustment sign-preserving techniques, including GRAS, by means of a numerical simulation. Finally, section five presents the main conclusions and finishes the paper.

2. Formulation

Consider a prior $(T \times K)$ matrix $\mathbf{A}$ with cells a_{ij} to be adjusted to a target matrix $\mathbf{X}$ with unknown cells x_{ij}, but with observable row and column totals $\boldsymbol{u}$ and $\boldsymbol{v}$ respectively. The traditional GRAS problem is to find the matrix $\mathbf{X}$ that deviates least from $\mathbf{A}$ and is consistent with the row and columns margins.

Junius and Oosterhaven (2003) formulated their proposed solution as a variant of the traditional RAS problem, but allowing for the presence of both positive and negative entries. Lenzen et al. (2007) suggested some modifications in the target function in order to account for the distance between the initial and the target matrix. The formulation proposed in Lenzen et al. (2007) is:

$$z_{ij} = \arg\ min \sum_{i=1}^{T}\sum_{j=1}^{K} |a_{ij}|\, z_{ij} ln\left(\frac{z_{ij}}{e}\right) \qquad \text{being } z_{ij} = x_{ij}/a_{ij} \qquad (1)$$

And e is the base of the natural logarithm. This minimization is subject to the row and column constrains:

$$\begin{aligned} \textstyle\sum_{i=1}^{T} a_{ij} z_{ij} = \sum_{i=1}^{T} x_{ij} = v_j; \\ \textstyle\sum_{j=1}^{K} a_{ij} z_{ij} = \sum_{j=1}^{K} x_{ij} = u_i \end{aligned} \qquad (2)$$

In this paper an alternative updating method applicable for matrices with both positive and negative cells is proposed. The proposed technique can be seen as an extension of the paper by Golan et al. (1994). In that paper, a Generalized Cross Entropy (GCE) procedure was proposed to recover intersectoral information from incomplete data. The context for applying this idea was, however, somewhat restricted since it only considered matrices of coefficients (bounded between 0 and 1). In this article this method is extended to cases where the entries of the target matrix are flows instead of coefficients and that can contain both positive (larger than 1) and negative cells.

The point of departure is considering each element of the prior and target matrices, **A** and **X**, as realizations of random variables that can take a range of M possible values which are contained in a vector $\boldsymbol{b'_{ij}} = \left[b_{ij1}, \dots, b^{*}_{ij}, \dots, b_{ijM}\right]$ with values that are set exogenously. Each support vector $\boldsymbol{b_{ij}}$ can be different

for every cell and contains an odd number of values that are centered on point b_{ij}^* symmetrically. The entries in the prior matrix determine the central points b_{ij}^* for each vector. More specifically, each cell a_{ij} in the initial matrix **A** is assumed to be this particular point of its corresponding vector ($a_{ij} = b_{ij}^*$), although any of the other points contained in $\boldsymbol{b_{ij}}$ could have been observed instead. These other values for each vector are specified arbitrarily by the researcher, depending on our beliefs about how much it is possible to deviate from b_{ij}^*.

For the sake of simplicity, let us illustrate this idea by considering the simplest case whit $M = 3$. In this situation, the support vector would be defined as $\boldsymbol{b'_{ij}} = \left[(1-r)a_{ij}\,, a_{ij}\,, (1+r)a_{ij}\right] = \left[b_{ij1}\,, b_{ij}^*\,, b_{ij2}\right]$. The scalar r represents a rate of variation imposed by the researcher with respect to a_{ij}, which determines the minimum and maximum value assumed as possible for this cell. Note that if we set any $|r| \leq 1$, we prevent the possibility that this element could change its sign from positive to negative or vice versa, but this sign-preserving character can be removed just by setting a scalar $|r| > 1$.[4]

Once the possible realizations for each entry in the matrices have been specified, given that we assume that they are generated by a random process, some probability distribution should be assigned to them. Although the support vectors for the cells of **A** and **X** are common, the distribution probabilities are different. In the case of **A**, these probabilities are set a priori by the researcher, but they are unknown for our target matrix **X**.

Starting with the elements of matrix **A**, we need to specify a probability distribution as $\boldsymbol{q'_{ij}} = \left[q_{ij1}, \ldots, q_{ij}^*, \ldots, q_{ijM}\right]$ for each element a_{ij}. Continuing with the simplest case with $M = 3$, one natural way of doing this is by

[4] I assume here that this scalar is common to all the cells in the matrix, but this assumption can be relaxed easily.

assuming that all the values are equally probable and setting $q_{ij1} = q_{ij}^* = q_{ij2} = 1/3$. This solution gives to the value actually observed (b_{ij}^*) the same probability as to the extreme cases b_{ij1} and b_{ij2}. An alternative could be to assign an arbitrarily high probability q_{ij}^* to b_{ij}^* and to assume that the two extreme cases are equally probable to each other. Whatever the specific probabilities chosen, the general rule $q_{ij1} = q_{ij2} = (1 - q_{ij}^*)/2$ guarantees that:

$$a_{ij} = \sum_{m=1}^{M} q_{ijm} b_{ijm} \tag{3}$$

We apply the same reasoning with the elements of the target matrix **X**, but now the probability distributions $\boldsymbol{p'_{ij}} = [p_{ij1}, p_{ij2}, \dots, p_{ijM}]$ are unknown and must be estimated. The value of each cell of this matrix is given by the expression:

$$x_{ij} = \sum_{m=1}^{M} p_{ijm} b_{ijm} \tag{4}$$

In this framework of analysis, the original problem of adjusting matrix **X** from matrix **A**, has been transformed in a new problem where a set of posterior probabilities **P** will be estimated from the a priori probabilities **Q**.

3. The GCE solution

3.1. The GCE procedure: numerical optimization

A constrained minimization problem is applied in order to find the solution to the GCE estimator. The estimation problem can be posed as a minimization program like:

$$\underset{\boldsymbol{P}}{Min}\, D(\mathbf{P}\|\mathbf{Q}) = \sum_{m=1}^{M}\sum_{i=1}^{T}\sum_{j=1}^{K} p_{ijm}\, ln\left(\frac{p_{ijm}}{q_{ijm}}\right) \tag{5}$$

Subject to:

$$\sum_{j=1}^{K} x_{ij} = \sum_{j=1}^{K}\left(\sum_{m=1}^{M} p_{ijm} b_{ijm}\right) = v_i\,;\ \forall i \tag{6}$$

$$\sum_{i=1}^{T} x_{ij} = \sum_{i=1}^{T}\left(\sum_{m=1}^{M} p_{ijm} b_{ijm}\right) = u_j\,;\ \forall j \tag{7}$$

In the original paper by Junius and Oosterhaven (2003, pp. 90-91) and in the correction proposed by Lenzen et al. (2007, pp. 464-465) proofs of the bi-proportionality of the solution of the GRAS algorithm are presented. In a similar fashion, this section presents the solution of the GCE program contained in equations (5)-(7) and it shows that the solution –estimates of the target matrix- achieved are not biproportional to the information contained in the prior matrix.

The Lagrangean function related to (5)-(7) is:

$$\begin{aligned}\mathcal{L} = \sum_{m=1}^{M}\sum_{i=1}^{T}\sum_{j=1}^{K} p_{ijm}\, ln\left(\frac{p_{ijm}}{q_{ijm}}\right) + \sum_{j=1}^{K}\lambda_j\left[u_j - \sum_{i=1}^{T}\left(\sum_{m=1}^{M} p_{ijm} b_{ijm}\right)\right] \\ + \sum_{i=1}^{T}\pi_i\left[v_i - \sum_{j=1}^{K}\left(\sum_{m=1}^{M} p_{ijm} b_{ijm}\right)\right]\end{aligned} \tag{8}$$

With corresponding derivatives:

$$\frac{\partial \mathcal{L}}{\partial p_{ijm}} = ln\left(\frac{p_{ijm}}{q_{ijm}}\right) + 1 - \lambda_j b_{ijm} - \pi_i b_{ijm} = 0 \tag{9}$$

Imposing the optimality conditions in (9) yields:

$$ln(p_{ijm}) = \pi_i b_{ijm} + ln(q_{ijm}) + (\lambda_j b_{ijm} - 1);$$
$$\text{or} \qquad (10)$$
$$p_{ijm} = exp(\pi_i b_{ijm}) q_{ijm} exp(\lambda_j b_{ijm} - 1) = \tilde{\rho}_{ijm} q_{ijm} \tilde{\sigma}_{ijm}$$

Being $\tilde{\rho}_{ijm} = exp(\pi_i b_{ijm})$ and $\tilde{\sigma}_{ijm} = exp(\lambda_j b_{ijm} - 1)$. Note that this biproportional relationship between the a priori and posterior distributions **Q** and **P** does not hold for the prior and target matrices **A** and **X**. This means that the GCE solution is not necessarily sign-preserving, but depends on the absolute value of scalar r used to set the possible values included in the support vector.

3.2. Inference

One of the main advantages of the proposed GCE procedure is that, contrary to biproportional techniques like GRAS, makes possible doing some inference with the estimates, following Golan et al. (1994). Once the optimization problem depicted in equations (5) to (7) is solved and the elements $\hat{p}_{ijm}$ are recovered, point estimates of the entries of **X** are calculated as in (4) by the expression $\hat{x}_{ij} = \sum_{m=1}^{M} \hat{p}_{ijm} b_{ijm}$. These point estimates can be complemented by calculating some indicator of the variability in $\hat{x}_{ij}$, given that the stochastics nature of the elements x_{ij} present in the GCE procedure allows for estimating their variances as well. The following expression:

$$Var(\hat{x}_{ij}) = \sum_{m=1}^{M} [\hat{p}_{ijm} b_{ijm}^2] - \left[\sum_{m=1}^{M} \hat{p}_{ijm} b_{ijm}\right]^2 \qquad (11)$$

calculates the estimated variances of each estimate in the target matrix. An additional indicator of uncertainty for each estimated cell through the Shannon's entropy measure:

$$S(\hat{x}_{ij}) = -\sum_{m=1}^{M} \hat{p}_{ijm}\, ln(\hat{p}_{ijm}), \tag{12}$$

which can be conveniently re-scaled from 0 to 1 by dividing it by its maximum value $ln(M)$. Values of $S(\hat{x}_{ij})/ln(M)$ close to 1 will be indicating a high degree of uncertainty associated to the estimate of entry x_{ij}, while the opposite situation happens for values of $S(\hat{x}_{ij})/ln(M)$ close to 0.

Additionally, hypothesis testing is also possible in the GCE framework basing on the relationship between the objective functions of restricted and unrestricted GCE problems. Let $D_U(\mathbf{P}\|\mathbf{Q}) = \sum_{m=1}^{M}\sum_{i=1}^{T}\sum_{j=1}^{K} \hat{p}_{ijm}\, ln\left(\frac{\hat{p}_{ijm}}{q_{ijm}}\right)$ be the Kullback-Leibler divergence evaluated at the solutions $\hat{p}_{ijm}$ of optimization problem as in equations (5) to (7) and $D_R(\mathbf{P}\|\mathbf{Q})$ be the same function where the solutions are restricted to fulfil J additional constraint ($\hat{x}_{ij} = 0$, for example, if $J = 1$). Under some mild assumptions, Golan et al. (2000, pages 407-408) show that:

$$2[D_R(\mathbf{P}\|\mathbf{Q}) - D_U(\mathbf{P}\|\mathbf{Q})] \rightarrow \chi^2_J \tag{13}$$

3.3. An illustrative example

This subsection illustrates how the GCE procedure can be implemented by solving the same balancing problem as in Junius and Oosterhaven (2003). As point of departure, the same initial matrix $\boldsymbol{A}$ used by Junius and Oosterhaven (2003, page 94) to illustrate the GRAS procedure is taken as reference:

<<Insert

Table 1 around here>>

Cells in Table 1 will be adjusted to make them consistent with the totals observable for the posterior matrix, as presented in Table 2:

<<Insert Table 2 around here>>

Applying GCE requires the specification of the M points contained in the support vectors $\boldsymbol{b'_{ij}}$ and that define the possible values taken in the target cells x_{ij}. I opted for a simple case with $M = 3$, where $\boldsymbol{b'_{ij}} = [(1-r)a_{ij}, a_{ij}, (1+r)a_{ij}]$ and setting $r = 0.5$, which allows each cell in the initial matrix to be adjusted to $\pm 50\%$ of its value as maximum.[5]

The probability distributions $\boldsymbol{q'_{ij}}$ associated to each element a_{ij} are the other important point in the GCE adjustment. They implicitly reflect our beliefs about how much deviation can be assumed between the observed realization in the cell a_{ij} and its unknown counterpart x_{ij}. If we believe that the "extreme" values $(1-r)a_{ij}$ or $(1+r)a_{ij}$ are not probable -i.e., the x_{ij} element are expected to be close to the initial cell a_{ij}-, we can assign a prior distribution with a mass probability in the central point and $q_{ijm} \simeq 0$ for the rest of values. If, on the contrary, we consider that the entry x_{ij} is not necessarily very close to the initial a_{ij} but it can take values across all the parameter space defined in $\boldsymbol{b'_{ij}}$ with equal probability, we can assume an uniform distribution $q_{ij1} = q^*_{ij} = q_{ij2} = 1/3$. In this example I apply this uniform distribution and also a "spike" one as $\boldsymbol{q'_{ij}} = [0.025, 0.95, 0.025]$.

Table 3 and Table 4 present the solutions produced by GCE under the two alternative a priori probability distributions considered, uniform and spike respectively. Each cell in these tables report the point estimates of each cell,

[5] For the sake of simplicity, in this example I prevent sign flips by setting a scalar $r < 1$.

together with estimations of its standard deviation –in brackets- and the value of the normalized entropy indicator –in parentheses- described in (12).

<<Insert Table 3 around here>>

<<Insert Table 4 around here>>

The solutions reported in these two tables are similar to each other in terms of the point estimates- and they are, in turn, similar to the GRAS solution in the original paper by Junius and Oosterhaven- indicating that the choice of the a priori probability distribution does not condition largely the estimates in this example. There are bigger differences, however, regarding the variability of the estimates, both in terms of their variance and the indicator of uncertainty measured by the normalized Shanon's entropy indicator, which are lower in general when a spike a priori distribution is chosen. Despite the differences between both cases, the results show how some particular cells – like the Net exports of 'Goods' or the Final Consumption of 'Services', for example- exhibit large variability and uncertainty indicators. This outcome is an indication that the estimation of these cells can be problematic in terms of their reliability and that collection of data regarding these cases is particularly important in order to alleviate these problems.[6]

4. A numerical experiment

In this section the proposed GCE solution will be compared with the other techniques by means of a numerical simulation under several possible scenarios. Again, the same initial matrix **A** shown in Table 1 will be the point of departure for the experiment. The target matrix to be estimated is different

[6] Large variability or uncertainty indicators can be taken as a signal of higher risk of making large estimation errors in these specific cells. See Hosoe (20014) for a recent study on the consequence of errors when estimating IO datasets that are later used as the basis for CGE models.

in in each trial of the numerical simulations and it is generated modifying each cell of Table 1 by introducing some noise:

$$x_{ij} = a_{ij} \times \varepsilon_{ij}, \text{ where } \varepsilon_{ij} \sim N(1, \sigma) \tag{14}$$

This generation process for the cells of the target preserves the zeros present in the initial matrix. Additionally, the standard deviation σ conditions the distance between the initial and the final elements. Initially I set $\sigma = 0.1$, given that with such a standard deviation the possibility of changes in the sign of the cells is virtually prevented. Additionally I try with different values of σ (specifically $\sigma = 0.2$ and 0.5) in order to consider larger differences between **A** and **X**. Note that a standard deviation as 0.1 or 0.2 virtually prevents a change in the sign of the cell, but a standard deviation in ε_{ij} as big as 0.5 allows the possibility of such a change.[7] The row and column totals are assumed as observable in the target matrices generated and they are incorporated as constrains to the adjustment problem.

Again, the M points of the support vectors $\boldsymbol{b'_{ij}}$ are set for a case with $M = 3$, and the rate of maximum and minimum change that each initial value is assume to vary is given by scalar r. Specifically, the values $r = 1, 2$ and 10 are set in the experiment in order to have some indication about the sensitivity of the estimates to this specification. The a priori distributions $\boldsymbol{q'_{ij}}$ are also specified following the same approach as in the previous section and for each cell of the matrix, the GCE solutions have been obtained from two alternative a priori distributions: one "spike" distribution that assigns a probability close to one to the central values in the supporting vectors –and, consequently, a probability close to zero to values on the extremes- and one uniform

[7] Given the nature of the table taken as basis for the experiment, some of the cells cannot be negative by definition, so their sing in the target matrix should be preserved as positive. In order to keep the realism in the simulation, the cells reflecting flows from 'Goods' and 'Services' to themselves and to 'Consumption' in the simulated final matrices are generated as in equation (17), but replacing ε_{ij} by $|\varepsilon_{ij}|$ if it is originally generated negative. The deviation measures calculated in the experiment are not sensitive to this correction.

distribution with prior probability 1/3 for all the points in the supporting vectors.

The GCE solutions found in the experiment are compared with those by other balancing techniques, being one of them the GRAS algorithm. In order to enhance this comparison, other procedures have been considered as well. The recent papers by Huang et al. (2008), Pavia et al. (2009) or Termushoev et al. (2010) evaluated alternative adjustment techniques to the GRAS objective function (1), suggesting the three following variants:

$$z_{ij} = \arg\, min \sum_{i=1}^{T}\sum_{j=1}^{K} |a_{ij}|\left(z_{ij}-1\right)^2 \qquad \textit{Improved Normalized Squared Differences (INSD)} \tag{15}$$

$$z_{ij} = \arg\, min \sum_{i=1}^{T}\sum_{j=1}^{K} \left(a_{ij}\right)^2\left(z_{ij}-1\right)^2 \qquad \textit{Improved Squared Differences (ISD)} \tag{16}$$

$$z_{ij} = \arg\, min \sum_{i=1}^{T}\sum_{j=1}^{K} |a_{ij}^3|\left(z_{ij}-1\right)^2 \qquad \textit{Improved Weighted Squared Differences (IWSD)} \tag{17}$$

To evaluate the performance of these five estimation approaches (GCE, GRAS, INSD, ISD and IWSD), 1,000 trials have been carried out. There are several different deviation measures that can be applied to evaluate the adjustment (see Lahr 2001, appendix 3, for a survey of the possible measures). In the experiment I opted for calculating the *Weighted Absolute Percentage Error* (WAPE), which has been largely used when evaluating the performance of adjusting techniques (see Jiang et al., 2010a and 2010b, for recent examples). This measure averages the percentage error giving larger weights to errors in large cells than errors in small cells (Oosterhaven et al, 2008). It is defined as:

$$WAPE = \sum_{i=1}^{T}\sum_{j=1}^{K} 100\frac{\left|x_{ij}-\hat{x}_{ij}\right|}{\sum_{i=1}^{T}\sum_{j=1}^{K}\left|x_{ij}\right|} \tag{18}$$

where the $\hat{x}_{ij}$ elements denote the estimated entries. Additionally, the so-called *Standardized weighted absolute difference (SWAD)* is calculated as follows:

$$SWAD = \sum_{i=1}^{T}\sum_{j=1}^{K} \frac{|x_{ij}| \times |x_{ij} - \hat{x}_{ij}|}{\sum_{i=1}^{T}\sum_{j=1}^{K}[x_{ij}]^2} \tag{19}$$

The SWAD is a deviation measure similar to WAPE, but now the absolute deviations are weighted by the size of the true transactions (Lahr, 2001). Table 5 shows the results.

<<Insert Table 5 around here>>

Deviation measures in Table 5 indicate a very similar performance between GRAS and INSD, which both clearly beat ISD and IWSD under any of the three scenarios simulated. These results are similar to those reported in Temurshoev et al. (2011) where several adjusting techniques were evaluated by means of an empirical application for The Netherlands and Spain (see Tables 2, 3 and 4 in Temurshoev et al., 2011). The proposed GCE technique, however, slightly outperforms GRAS and INSD and the gains in comparatively smaller deviations become larger when scalar σ grows. This result is not surprising, given that the GCE technique is not a strictly sign-preserving: it departs from the cell present in the prior matrix but allows for a potential change of sign in the corresponding posterior cell. We can assume this change as more or less likely by setting the a priori probabilities $\boldsymbol{q}$. Generally speaking, the higher the probability assigned to the central point in the support vectors (q_{ij}^{*}), the smaller the probability of a change in the sign of the solution. The performance of the technique seems relatively insensitive to changes in the support vectors (by changing the scalar r) or to changes in the a priori distributions set in $\boldsymbol{q}$.

5. Concluding remarks

An adjustment technique for matrices with positive and negative cells has been proposed in this paper. The suggested Generalized Cross-Entropy (GCE) method has as one its main advantages a higher flexibility when compared with other traditional sign-preserving techniques. Given that it requires the specification of a supporting vector containing all the possible realizations of each cell, it allows for preventing changes in the sign simply by not considering values that could lead to such a change. Alternatively, supporting vectors with values that change the sign of a cell can be included with an arbitrarily low a priori probability. This situation can reflect researcher's belief about the behavior of a specific entry in a matrix, where a change in its sign can be improbable but not totally impossible. Additionally, GCE offers the possibility of doing some inference with the estimates, something that is not possible with traditional balancing techniques based on biproportional adjustment.

The numerical experiment conducted in the paper, producing smaller deviation measures than other competing procedures, suggests that the proposed GCE technique can be considered as an alternative to other adjustment methods. The possibility of deriving indicators of variability and uncertainty for each estimated cell of the matrix is also attractive, since it allows for identifying specific cells that can be problematic. Interestingly, one potential application of this technique is to use it combined with other sign-preserving techniques like GRAS: in the estimation of Supply and Use Tables (SUT's), GRAS can be used when changes in the sign are virtually impossible –i.e., in the case of the matrix of intermediate demand- while GCE can be applied only to parts of the table where these changes are possible –i.e., the final demand matrix-.

References

Golan A, Judge G, Robinson S, 1994, "Recovering information from incomplete or partial multisectoral economic data", *Review of Economics and Statistics*, 76, 541-549

Golan A, Karpf LS, Perloff J, 2000, "Estimating Coke's and Pepsi's price and advertising strategies", *Journal of Business and Economic Statistics*, 18, 398-409

Günlük-Senesen G, JM Bates, 1988, "Some experiments with methods of adjusting unbalanced data matrices", *Journal of the Royal Statistical Society, Series A*, 151, 473–490

Hosoe N, 2014, "Estimation errors in input–output tables and prediction errors in computable general equilibrium analysis", *Economic Modelling*, 42, 277–286

Huang W, Kobayashi S, Tanji H, 2008, "Updating an Input–Output matrix with sign-preservation: some improved objective functions and their solutions", *Economic Systems Research*, 20, 111–123

Jiang X, Dietzenbacher E, Los B, 2010a, "Targeting the collection of superior data for the estimation of the intermediate deliveries in regional input - output tables", *Environment and Planning A*, 42, 2508-2526

Jiang X, Dietzenbacher E, Los B, 2010b, "Improved estimation of regional Input–Output tables using cross-regional methods", *Regional Studies*, 465, 1-17

Junius T, Osterhaven J, 2003, "The solution of updating or regionalizing a matrix with both positive and negative entries", *Economic Systems Research*, 15, 87–96

Lahr ML, 2001, "A Strategy for producing hybrid regional input–output tables", in *Input–Output Analysis: Frontiers and Extensions*, Eds ML Lahr, E Dietzenbacher, (Palgrave, New York) pp 211–242

Lenzen, M, Wood R, Gallego B, 2007, "Some comments on the GRAS method", *Economic Systems Research*, 19, 461–465

Lenzen, M, Moran DD, Geschke A, Kanemoto K, 2014, "A non-sign-preserving RAS variant", *Economic Systems Research*, 26, 197-208

Miller RE, Blair PD, 2009, *Input-output analysis: Foundations and extensions*, Cambridge University Press, Cambridge

Oosterhaven, J, Stelder D, Inomata S, 2008, "Estimating international interindustry linkages: non-survey simulations of the Asian-Pacific economy", *Economic Systems Research*, 20, 395-414

Pavia JM, Cabrer B, Sala R, 2009, "Updating input–output matrices: assessing alternatives through simulation", *Journal of Statistical Computation and Simulation*, 79, 1467-1482

Temurshoev U, Webb C, Yamano N, 2011, "Projection of supply and use tables: Methods and their empirical assessment", *Economic Systems Research*, 23, 91–123

Temurshoev U, Miller RE, Bouwmeester M, 2013, "A note on the GRAS method", *Economic Systems Research*, 25, 361–367

Table 1. Initial matrix to be adjusted

	Goods	Services	Consumption	Net exports	Total output
Goods	7	3	5	-3	**12**
Services	2	9	8	1	**20**
Net Taxes	-2	0	2	1	**1**
Total Use	**7**	**12**	**15**	**-1**	**33**
Value added	**5**	**8**	**0**	**0**	**13**
Total input	**12**	**20**	**15**	**-1**	

Source: Junius and Oosterhaven (2003, page 94)

Table 2. Target matrix, only row and column totals observable

	Goods	Services	Consumption	Net exports	Total output
Goods					**15**
Services					**26**
Net Taxes					**-1**
Total Use	**9**	**16**	**17**	**-2**	**40**
Value added	**6**	**10**	**0**	**0**	**16**
Total input	**15**	**26**	**17**	**-2**	

Source: Junius and Oosterhaven (2003, page 94)

Table 3. GCE solution, uniform a priori distribution

	Goods	Services	Consumption	Net exports	Total output
	9.51	3.44	5.66	-3.60	
Goods	[1.91]	[1.16]	[1.96]	[1.11]	**15**
	(0.61)	(0.94)	(0.95)	(0.89)	
	2.30	12.56	10.19	0.94	
Services	[0.77]	[2.12]	[1.96]	[0.41]	**26**
	(0.94)	(0.50)	(0.95)	(0.99)	
	-2.81		1.15	0.66	
Net Taxes	[0.45]		[0.40]	[0.29]	**-1**
	(0.48)		(0.40)	(0.64)	
Total Use	**9**	**16**	**17**	**-2**	**40**
Value added	**6**	**10**	**0**	**0**	**16**
Total input	**15**	**26**	**17**	**-2**	

Note: point estimates are reported on each cell, together with its standard deviation in brackets and the normalized entropy indicator $S(\hat{x}_{ij})/ln(M)$ in parentheses.

Table 4. GCE solution, "spike" a priori distribution

	Goods	Services	Consumption	Net exports	Total output
	9.73	3.22	5.68	-3.63	
Goods	[1.45]	[0.55]	[1.12]	[0.74]	**15**
	(0.48)	(0.41)	(0.54)	(0.62)	
	2.08	12.78	10.18	0.97	
Services	[0.29]	[1.65]	[1.99]	[0.14]	**26**
	(0.30)	(0.40)	(0.63)	(0.29)	
	-2.81		1.14	0.66	
Net Taxes	[0.39]		[0.35]	[0.24]	**-1**
	(0.44)		(0.38)	(0.58)	
Total Use	**9**	**16**	**17**	**-2**	**40**
Value added	**6**	**10**	**0**	**0**	**16**
Total input	**15**	**26**	**17**	**-2**	

Note: point estimates are reported on each cell, together with its standard deviation in brackets and the normalized entropy indicator $S(\hat{x}_{ij})/ln(M)$ in parentheses.

Table 5. Deviations between target and estimates in the numerical simulation (1,000 trials)

	$\varepsilon_{ij} \sim N(1,0.1)$		
	Technique	**WAPE (%)**	**SWAD**
	GRAS	4.03	0.004
	INSD	4.04	0.004
	ISD	5.96	0.006
	IWSD	9.42	0.009
$\boldsymbol{q'}_{ij} = [0.333, 0.333, 0.333]$	GCE ($r = 1$)	3.61	0.004
	GCE ($r = 2$)	3.61	0.004
	GCE ($r = 10$)	3.61	0.004
$\boldsymbol{q'}_{ij} = [0.025, 0.95, 0.025]$	GCE ($r = 1$)	3.60	0.004
	GCE ($r = 2$)	3.60	0.004
	GCE ($r = 10$)	3.61	0.004
	$\varepsilon_{ij} \sim N(1,0.2)$		
	Technique	**WAPE (%)**	**SWAD**
	GRAS	8.09	0.009
	INSD	8.10	0.009
	ISD	11.92	0.012
	IWSD	18.56	0.018
$\boldsymbol{q'}_{ij} = [0.333, 0.333, 0.333]$	GCE ($r = 1$)	7.22	0.008
	GCE ($r = 2$)	7.22	0.008
	GCE ($r = 10$)	7.22	0.008
$\boldsymbol{q'}_{ij} = [0.025, 0.95, 0.025]$	GCE ($r = 1$)	7.21	0.008
	GCE ($r = 2$)	7.20	0.008
	GCE ($r = 10$)	7.21	0.008
	$\varepsilon_{ij} \sim N(1,0.5)$		
	Technique	**WAPE (%)**	**SWAD**
	GRAS	21.34	0.023
	INSD	20.69	0.022
	ISD	29.24	0.031
	IWSD	40.54	0.042
$\boldsymbol{q'}_{ij} = [0.333, 0.333, 0.333]$	GCE ($r = 1$)	18.12	0.019
	GCE ($r = 2$)	17.93	0.019
	GCE ($r = 10$)	18.43	0.020
$\boldsymbol{q'}_{ij} = [0.025, 0.95, 0.025]$	GCE ($r = 1$)	18.26	0.020
	GCE ($r = 2$)	18.22	0.019
	GCE ($r = 10$)	18.00	0.019