



**Projection of Supply and Use tables: methods  
and their empirical assessment**

Working Paper Number: 2

Authors: Umed Temurshoev, Norihiko Yamano and Colin Webb

# Projection of Supply and Use tables: methods and their empirical assessment\*

Umed Temurshoev<sup>†</sup>    Norihiko Yamano<sup>‡</sup>    Colin Webb<sup>‡</sup>

## Abstract

In this paper we present and perform the relative performance test of eight existing updating methods for the projection of Supply and Use tables (SUTs) in the example of the Netherlands and Spain. Some of the presented methods are less (or not) known in the literature, and have been slightly revised in projecting negative elements and preserving the signs of the entries of original matrices in the estimated ones. It has been found that (G)RAS and two other methods proposed, respectively, by Harthoorn and van Dalen (1987) and Kuroda (1988) produce the best estimates for the data at hand. Their relative success also suggests the stability of transactions' ratios for larger deliveries.

**Keywords:** Supply, Use, updating techniques

**JEL Classification Codes:** C61, C89, D57

---

\*We gratefully acknowledge financial support provided by EU FP7 WIOD project. This project is funded by the European Commission, Research Directorate General as part of the 7th Framework Programme, Theme 8: Socio-Economic Sciences and Humanities. Grant Agreement no: 225 281, [www.wiod.org](http://www.wiod.org). This paper was partly written when the first author was in a short visit to the Organisation for Economic Co-operation and Development (OECD) in Paris. He thanks the staff of the Directorate for Science, Technology and Industry of the OECD for their hospitality. We also thank the seminars' participants in the OECD and University of Groningen for their comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily represent those of the OECD.

<sup>†</sup>Department of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. E-mail: [u.temurshoev@rug.nl](mailto:u.temurshoev@rug.nl) (corresponding author)

<sup>‡</sup>Directorate for Science, Technology and Industry (DSTI), Organisation for Economic Co-operation and Development (OECD). E-mail: [yamano.norihiko@gmail.com](mailto:yamano.norihiko@gmail.com) and [colin.webb@oecd.org](mailto:colin.webb@oecd.org)

# 1 Introduction

It is well-known that Supply and Use tables (SUTs) are the core of the System of National Accounts, and provide the detailed picture of the entire economy. These tables provide a thorough overview of the production process, of the use of goods and services (commodities), and of income generated in the production process. SUTs are building blocks of the the so-called symmetric input-output tables (SIOTs), which together provide a powerful tool for very different types of economic analyses. The areas that make extensive use of these datasets are, for example, environmental analysis, globalization, international trade, productivity and efficiency, innovation and R&D, and migration.

Although the delivery and quality of SUTs provided by different statistics committees around the world are much better now than 10-15 years ago, still the problem of timeliness of these data remains. This has to do with the large financial expenditures and human efforts required in collecting SUTs. That is why most statistical committees of countries provide benchmark tables based on detailed surveys on usually five years interval. However, for the policy research objectives these five-year tables are often not adequate, hence the in-between benchmark SUTs and SIOTs are estimated. For this purpose, very different *non-survey* methods have been proposed and extensively used.

The purpose of this paper is twofold. First, we present eight methods of updating SUTs, four of which are less (or not) known in the literature. These less known updating techniques are the methods of EUKLEMS (Timmer et al. 2005), Euro (Beutel 2002, Eurostat 2008),<sup>1</sup> Harthoorn and van Dalen (1987), and Kuroda (1988). Other considered approaches are the well-known (G)RAS (see e.g., Leontief 1941, Stone 1961, Bacharach 1970), and the family of least squares ‘distance’ functions-based mathematical programs: improved normalized squared differences (Friedlander 1961, Huang et al. 2008), improved squared differences (Almon 1968, Jackson and Murray 2004, Huang et al. 2008), and improved weighted squared differences. Second, we empirically assess the relative performance of all the above mentioned updating methods in the example of Dutch and Spanish SUTs. We should note that all these methods are equally applicable for updating SIOTs as well. To the best of our knowledge, this is the first attempt to perform the relative performance test of the mentioned less known updating methods for SUTs estimation. The EUKLEMS method is discussed fully for the first time in this work. Further, Harthoorn and van Dalen’s (1987) and Kuroda’s (1988) methods, also for

---

<sup>1</sup>The Euro method was originally devised for updating symmetric input-output tables, but is also used in a SUT-setting, see a report prepared by Joerg Beutel to the European Commission (e.g., contract number 1508302007 FISC-D, April 2008). The method will be presented in detail in this paper.

some reason less known and rarely used in the related literature, were slightly revised in this paper with respect to (i) considering negative elements and (ii) preserving signs of the updated tables' elements. The well-known RAS method is a biproportional technique that estimates a new matrix from an initial matrix by scaling row- and column-wise the entries in the initial matrix such that it satisfies pre-specified row and column totals of the projection table.<sup>2</sup> RAS can handle only non-negative matrices, but in many cases one needs to estimate a new matrix that may have negative entries as well, which is the case for SUTs updating. Generalized RAS (GRAS) due to Günlük-Şenesen and Bates (1988) and Junius and Oosterhaven (2003) allows also for updating matrices with negative entries, hence as a generalization of RAS will be discussed in the next section.

It is impossible to consider *all* updating methods, because theoretically their number is infinity. For example, consider an estimation technique that is based on a (linear or nonlinear) program, which searches for the minimum 'distance' between the original and a new matrix subject to certain constraints. But the definition of 'distance' is rather arbitrary, and one can define an infinite number of functions that can be used as a measure of distance. However, there are some updating methods that we deliberately do not discuss here, and the reasons are explained in what follows. The problem of the minimum sum of *cross entropies* (CE) (Golan et al. 1994, Golan and Vogel 2000) is not considered for two reasons. First, it is applicable only to semipositive matrices, but in SUTs one can always find negative entries (for example, in the valuation adjustment and final demand matrices). Second, its objective function is biased (see for this point Lenzen et al. 2007): CE solutions tend to be  $1/e$  of the original elements, where  $e$  is the base of natural logarithm; one instead wants them to be as close as possible to the 'base' year table.<sup>3</sup> The TAU and UAT methods proposed by Snower (1990) are not applicable for updating SUTs as they are based on the Leontief quantity and price models, hence are (only) valid for estimating SIOTs. One, of course, can transform the updated SIOTs into the SUTs (on the base of the so-called product or industry technology assumption), but then a lot of information is lost, and, to our view, leads to poor estimates of SUTs. Further, it requires substantially more information than the eight methods analyzed in this paper. We do not consider a new variant of RAS – the so-called Cell-corrected RAS (CRAS) proposed recently by Mínguez et al. (2009). The reasons are as follows. First, CRAS at this moment can handle only non-negative matrices, and, second, it requires the availability of rather long time-series of SIOTs, which is not

---

<sup>2</sup>See Lahr and de Mesnard (2004) for details on RAS (including its history), which also gives an extensive set of references on the topic.

<sup>3</sup>See McDougall (1999) for detailed comparison of RAS and entropy-theoretic methods, who argues that, in general, the RAS remains preferable matrix balancing technique.

the case for many countries, especially, developing ones. Next, we did not analyze the so-called univariate updating methods, such as Proportional Correction Method (see e.g., Eurostat 2008, pp. 449-451) and Statistical Correction Method (Tilanus 1968). These methods ‘correct’ the original matrix only row-wise with a diagonal matrix of correction factors in order to get the updated matrix (hence, the term univariate). We think that such a correction is rather simple and less realistic than the RAS method, which corrects the original matrix both row-wise and column-wise to get the final estimate.<sup>4</sup> A recent updating method is a multi-objective approach (Strømman 2009), which combines two different programs by introducing one of the objective functions as a constraint into the other program. This approach needs a thorough analysis (e.g., dealing with negative elements), hence is not analyzed in this paper either.

The literature on evaluating non-survey methods is rather large. Some early work include Allen and Gossling (1975), Davis et al. (1977), Parikh (1979), Butterfield and Mules (1980), Günlük-Şenesen and Bates (1988) and St Louis (1989). We should mention that the RAS procedure was extensively used in other fields, such as demography (see e.g., Rees 1979, Schoen and Jonsson 2003), stochastic modeling (see e.g., Mari Bhat 1990), political studies (see e.g., Johnston and Pattie 1993), and transportation research (see e.g., Carey et al. 1981). The ‘novelty’ of this paper is that we also consider methods that, to our knowledge, have never been evaluated before in such a comparative setting, and implement the performance test on SUTs rather than SIOTs.

The rest of this paper is organized as follows. In Section 2 we present the above mentioned eight updating methods. Their empirical evaluation is carried out in Section 3 using the Dutch and Spanish SUTs. Some concluding remarks are given in Section 4.

## 2 Review of the updating methods

In this section we provide the description of different methods of updating Supply and Use tables (SUTs) and input-output (IO) tables. The attempt was to employ the notations used in the IO literature as much as possible. Adopting usual convention, matrices are given in bold, capital letters; vectors in bold, lower case letters; and scalars in italicized lower case letters. Vectors are columns by definition, thus row

---

<sup>4</sup>The three-stage RAS (TRAS) proposed by Gilchrist and St. Louis (1999) extends the original RAS by including additional information on the individual cells, and row and column totals of a submatrix of the original table. This variation of RAS is not considered here, since it is, in general, found that introduction of accurate exogenous information into RAS (besides row and column sums of the projection table) improves the resulting estimates (see e.g., de Mesnard and Miller 2006).

vectors are obtained by transposition, indicated by a prime.  $\hat{\mathbf{x}}$  denotes the  $n \times n$  diagonal matrix with the elements of the vector  $\mathbf{x}$  on its main diagonal and zero otherwise.

## 2.1 EUKLEMS method

In what follows we present in detail for the first time the method of estimating time series of SUTs in current prices that is partially discussed in Timmer et al. (2005), which we refer to as the EUKLEMS's method.

*Step 1: Data requirements.* For the benchmark year 0 we need the following data:

1.  $\mathbf{V}'_0$  – Supply matrix at basic price with dimension of commodity by industry (i.e.,  $com \times ind$ ). Note that its transpose  $\mathbf{V}$  is known as the Make matrix;
2.  $\mathbf{U}_0$  – Use matrix at purchase price (of dimension  $com \times ind$ );
3.  $\mathbf{Y}_0$  – Final demand matrix at purchase price ( $com \times f$ , where  $f$  is the number of final demand categories)
4.  $\mathbf{s}_0^{tr}$  – trade margin vector ( $com \times 1$ );
5.  $\mathbf{s}_0^{tt}$  – transportation margin vector ( $com \times 1$ );
6.  $\mathbf{s}_0^{vat}$  – VAT (non-deductable) vector ( $com \times 1$ );
7.  $\mathbf{s}_0^{nt}$  – the  $com \times 1$  vector of taxes on products net of subsidies.

For each projection period  $t$  the following data are needed:

8.  $\mathbf{x}_t$  – total industry output at basic price ( $ind \times 1$ );
9.  $\mathbf{u}_t$  – total industry intermediate inputs at purchase price ( $ind \times 1$ );
10.  $\mathbf{y}_t$  – total of final demand uses at purchase price ( $f \times 1$ );
11.  $m_t$  – total imports of all commodities (at cif) (only if imports by commodity is not available);
12.  $s_t^{tr}$  – total value of trade margins;
13.  $s_t^{tt}$  – total of transportation margins;
14.  $s_t^{vat}$  – total value of VAT (if available);
15.  $s_t^{nt}$  – total net taxes on products (if available).

Note that from the basic accounting equations it follows that  $\mathbf{x}_t = \mathbf{V}_t \mathbf{z}$  and  $\mathbf{u}_t = \mathbf{U}'_t \mathbf{z}$ , where  $\mathbf{z}$  is the summation vector of ones of appropriate dimension.

A general remark applicable to all methods considered in this paper is related to the use of *benchmark* SUTs in estimating new tables for some projection year(s). For *interpolation*, when data is available both for the beginning and end of a projection period, both benchmark data should be used. Assume we need to project SUTs for  $k$  years inbetween the start and ending periods with available SUTs. Then the benchmark data maybe taken to be simply as the arithmetic average of both available tables, for example,  $\mathbf{V}'_0 = (1/2)(\mathbf{V}'_{start} + \mathbf{V}'_{end})$ . However, since in some cases  $k$  can be large, it is better to use some weighting scheme that gives more weight to the

closer available data for the projection time  $t \in (0, k + 1)$ . Therefore, for example, the Supply matrix as the variable ‘benchmark’ depending on the projection year  $t$  can be written as  $\mathbf{V}'_0 = \frac{k+1-t}{k+1}\mathbf{V}'_{start} + \frac{t}{k+1}\mathbf{V}'_{end}$ . Similar variable benchmarks should be used for the base year Use tables and any other necessary benchmark table. The reason for such a weighting scheme is that structural change can be, at least, partially be taken into account by using variable benchmark tables. Note that if  $k = 1$  (i.e., projection of only one year SUTs is required), the base year data boils down to a simple arithmetic average of the two benchmarks as given earlier.

*Step 2: Generating Supply tables.* Compute the *industry output proportions* (or, equivalently, *product mix*) matrix for the base year as

$$\mathbf{C}_0 = \mathbf{V}'_0 \hat{\mathbf{x}}_0^{-1},$$

where  $\mathbf{x}_0 = \mathbf{V}_0 \mathbf{z}$ . Its typical element  $c_{ij}$  denotes the fraction of total industry  $j$ 's output that is in the form of commodity  $i$ . The EUKLEMS method assumes that the product mix matrix  $\mathbf{C}_0$  is valid for the entire projection period. Thus, the Supply matrix at time  $t$  is

$$\mathbf{V}'_t = \mathbf{C}_0 \hat{\mathbf{x}}_t. \quad (1)$$

Note that these estimated Supply matrices are consistent since  $\mathbf{z}'\mathbf{V}'_t = \mathbf{z}'\mathbf{C}_0\hat{\mathbf{x}}_t = \mathbf{z}'\mathbf{V}'_0\hat{\mathbf{x}}_0^{-1}\hat{\mathbf{x}}_t = \mathbf{x}'_t$ . If commodity imports vector is not available for the projection period, they are constructed as follows. Let  $\mathbf{m}_0$  be the vector of imports in the base year, then the projected imports vector is derived simply by applying the share of commodity imports compositions of the base year to the available aggregate imports value (data 11 in Step 1), i.e.,

$$\mathbf{m}_t = \frac{m_t}{\mathbf{z}'\mathbf{m}_0} \mathbf{m}_0.$$

Consequently, the time series of domestic commodity output and gross commodity output vectors at basic prices are found, respectively, as follows

$$\begin{aligned} \mathbf{q}_t^d &= \mathbf{V}'_t \mathbf{z}, \\ \mathbf{q}_t &= \mathbf{q}_t^d + \mathbf{m}_t. \end{aligned}$$

*Step 3: Generating gross commodity outputs at purchase prices.* The initial guess for the vectors of margins and taxes are constructed using their corresponding rates from the benchmark year. These rates are taken relative to the corresponding outputs at basic prices, i.e.,

$$\hat{\mathbf{r}}_0^k = \hat{\mathbf{s}}_0^k \hat{\mathbf{q}}_0^{-1} \quad \text{for } k = tr, tt, vat, nt,$$

where, for example,  $r_{i,0}^{tr} = s_{i,0}^{tr}/q_{i,0}^{tr}$ . The EUKLEMS method assumes that  $\hat{\mathbf{r}}_0$  is valid for the projection period, thus the initial guess, denoted by the subscript (1), of the trade and transportation margins, and VAT and other taxes is found as

$$\mathbf{s}_{t(1)}^k = \hat{\mathbf{r}}_0 \mathbf{q}_t \quad \text{for } k = tr, tt, vat, nt. \quad (2)$$

The totals of these estimated margins and taxes have to be consistent with their corresponding actual totals. Hence, given the time series of total margins and taxes (data 12-15 in step 1), the following normalization is used

$$\mathbf{s}_t^k = \frac{s_t^k}{\mathbf{1}' \mathbf{s}_{t(1)}^k} \mathbf{s}_{t(1)}^k \quad \text{for } k = tr, tt, vat, nt. \quad (3)$$

If the totals of taxes are not available, their initial estimates (2) are taken as the final estimates. The time series of the gross commodity outputs at purchase price (*pp*) is then generated as

$$\mathbf{q}_t^{pp} = \mathbf{q}_t + \mathbf{s}_t^{tr} + \mathbf{s}_t^{tt} + \mathbf{s}_t^{vat} + \mathbf{s}_t^{nt}.$$

*Step 4: Generating Use table at purchase price.* Simple correction method is used to equate supply and use of each commodity. The commodity growth rates are given by the diagonal matrix

$$\hat{\mathbf{g}}_t = \hat{\mathbf{q}}_t^{pp} (\hat{\mathbf{q}}_0^{pp})^{-1}.$$

Denote the  $com \times f$  final demand matrix by  $\mathbf{Y}$ . The initial guess of the Use table at time  $t$  is

$$[\mathbf{U}_{t(1)}, \mathbf{Y}_{t(1)}] = \hat{\mathbf{g}}_t [\mathbf{U}_0, \mathbf{Y}_0].$$

In general, it is the case that  $\mathbf{1}'[\mathbf{U}_{t(1)}, \mathbf{Y}_{t(1)}] \neq [\mathbf{u}'_t, \mathbf{y}'_t]$ , that is, the total estimated intermediate and final uses are not equal to their actual totals (data 9-10 in step 1). Thus, normalization is needed so that the estimated Use table becomes consistent with the totals of actual intermediate and final uses. This is implemented as follows

$$[\mathbf{U}_t, \mathbf{Y}_t] = [\mathbf{U}_{t(1)}, \mathbf{Y}_{t(1)}] \widehat{[\mathbf{u}_t, \mathbf{y}_t]} \widehat{[\mathbf{u}_{t(1)}, \mathbf{y}_{t(1)}]}^{-1}, \quad (4)$$

thus  $\mathbf{1}'[\mathbf{U}_t, \mathbf{Y}_t] = \mathbf{1}'[\mathbf{U}_{t(1)}, \mathbf{Y}_{t(1)}] \widehat{[\mathbf{u}_t, \mathbf{y}_t]} \widehat{[\mathbf{u}_{t(1)}, \mathbf{y}_{t(1)}]}^{-1} = [\mathbf{u}'_{t(1)}, \mathbf{y}'_{t(1)}] \widehat{[\mathbf{u}_{t(1)}, \mathbf{y}_{t(1)}]}^{-1} \widehat{[\mathbf{u}_t, \mathbf{y}_t]} = [\mathbf{u}'_t, \mathbf{y}'_t]$ .<sup>5</sup>

It is, however, very important to note that the final demand matrix excludes the column of changes in inventories and valuables, which is considered as a residual category in order to guarantee the consistency of the Supply and Use tables. That

---

<sup>5</sup>Note that steps (9)-(11) are similar to the first two steps in a RAS iteration procedure.



is, the vector of changes in inventories will be taken as the difference between total supply by product at purchase price derived in Step 3 and the sum of the intermediate and final uses by product (that excluded changes in inventories and valuables during the projection).

*Step 5: Generating Use table at basic price.* To transform the Use table at purchase price in (4) from Step 4 into basic prices, we need five valuation tables, namely, the matrices of: wholesale trade margins (WTR), retail trade margins (RTR), motor trade margins (MTR), transportation margins (TTM), and taxes net of subsidies margins (TXM). These matrices are found using the projected margins from (3). First, the following rates are defined:

$$\hat{\mathbf{g}}_t^k = \hat{\mathbf{s}}_t^k (\hat{\mathbf{q}}_t^{pp})^{-1} \quad \text{for } k = wtr, mtr, tt, tx \quad \text{and} \quad \hat{\mathbf{g}}_t^{rtr} = \hat{\mathbf{s}}_t^{rtr} \hat{\mathbf{f}}_t^{-1},$$

where  $\mathbf{f}_t = \mathbf{Y}_t \mathbf{t}$  is the final commodity demand vector at purchase prices. The reason for having separate rate for RTR is that retail trade is mainly used by the final demand categories, and only a minor part of it is due to the intermediate use. These rates are then applied row-wise to the projected Use tables at purchase prices in order to get the desired margins matrices. For example, the WTR and RTR matrices are estimated as

$$WTR_t = \hat{\mathbf{g}}_t^{wtr} [\mathbf{U}_t, \mathbf{Y}_t], \quad RTR_t = \hat{\mathbf{g}}_t^{rtr} \mathbf{Y}_t.$$

Other required matrices are estimated similar to the WTR projection. These are used to obtain the Use tables at basic prices as

$$[\mathbf{U}_t, \mathbf{Y}_t]^{bp} = [\mathbf{U}_t, \mathbf{Y}_t] - WTR_t - RTR_t - MTR_t - TTM_t - TXM_t. \quad (5)$$

Note that the resulting Use table (5) will be consistent with the Supply matrix at basic prices in (1), since the vectors of margins and taxes by product (3) are taken from the Supply table projection.

## 2.2 Euro method

The Euro method was developed by Joerg Beutel (see e.g., Eurostat 2008, Chapter 14).<sup>6</sup> The distinguishing feature of this method is that the estimated Supply and Use tables are based on macroeconomic forecasts of the growth rates of: (i) value added of industries, (ii) total final demand uses, and (iii) total imports. The method uses these official forecasts as exogenous input, and replicates them in the derived

---

<sup>6</sup>The SUT variant of the Euro method is given in different reports prepared by Joerg Beutel to the European Commission (e.g., contract number 150830-2007 FISC-D, April 2008).

SUTs, whose projection is based on two important assumptions. First, the shares of industries in the production of commodities remain constant, and second, the constant input coefficients (in each iteration) determine the relations of all commodity inputs to production of industries. In what follows we present the basics of the Euro method.

To implement the Euro method, it is required that the intermediate and final use tables are distinguished between the domestic and imported uses. Thus, the original base year SUTs consist of the following components all expressed at basic prices:

1.  $\mathbf{U}_0^d$  – domestic intermediate Use matrix ( $com \times ind$ ),
2.  $\mathbf{U}_0^m$  – imported intermediate Use matrix ( $com \times ind$ ),
3.  $\mathbf{Y}_0^d$  – domestic final demand matrix ( $com \times f$ ),
4.  $\mathbf{Y}_0^m$  – final demand matrix of imports ( $com \times f$ ),
5.  $\mathbf{V}_0$  – make matrix ( $ind \times com$ ),
6.  $\mathbf{v}_0$  – the vector of total value added of industries ( $ind \times 1$ ).

Further, the following macroeconomic forecasts for the projection year  $t$  are needed:

7.  $\mathbf{g}_t^y$  – growth rates of the value added of industries ( $ind \times 1$ ),
8.  $\mathbf{g}_t^f$  – growth rates of the totals of the final demand categories ( $f \times 1$ ), and
9.  $g_t^m$  – growth rate of total imports.

The data requirements for the projection year(s) are equivalent with the availability of the vectors of sectoral value-added,  $\mathbf{v}_t$ , of totals of final demand categories,  $\mathbf{y}_t$ , and aggregate value of imports,  $m_t$ .

It is obvious that the base year commodity outputs and imports vectors are equal to (recall that  $\mathbf{z}$  is the summation vector of ones of appropriate dimension)

$$\mathbf{q}_0^d = \mathbf{U}_0^d \mathbf{z} + \mathbf{Y}_0^d \mathbf{z} = \mathbf{V}_0' \mathbf{z}, \quad \text{and} \quad \mathbf{m}_0 = \mathbf{U}_0^m \mathbf{z} + \mathbf{Y}_0^m \mathbf{z}.$$

The *market share* (or, *commodity output proportions*) matrix of the base year is given by

$$\mathbf{D}_0 = \mathbf{V}_0 (\hat{\mathbf{q}}_0^d)^{-1},$$

whose typical element  $d_{ij}$  denotes the fraction of total commodity  $j$  output that is produced by industry  $i$ . It is assumed that the market share matrix is constant, thus for all industries and products in the projection years it is assumed that the share of industry  $i$  in the production of commodity  $j$  remains fixed.

Each iteration of the Euro method consists of two steps. The first step of the first iteration defines domestic and imported intermediate and final uses, the vector

of value added, and the make matrix of the projection year  $t$ , respectively, as follows

$$\begin{aligned}
\mathbf{U}_{t(1)}^d &= 0.5 \times (\hat{\mathbf{g}}_t^y \mathbf{U}_0^d + \mathbf{U}_0^d \hat{\mathbf{g}}_t^y), \\
\mathbf{U}_{t(1)}^m &= 0.5 \times (\hat{\mathbf{g}}_t^y \mathbf{U}_0^m + \mathbf{U}_0^m \hat{\mathbf{g}}_t^y), \\
\mathbf{Y}_{t(1)}^d &= 0.5 \times (\hat{\mathbf{g}}_t^y \mathbf{Y}_0^d + \mathbf{Y}_0^d \hat{\mathbf{g}}_t^y), \\
\mathbf{Y}_{t(1)}^m &= 0.5 \times (\hat{\mathbf{g}}_t^y \mathbf{Y}_0^m + \mathbf{Y}_0^m \hat{\mathbf{g}}_t^y), \\
\mathbf{v}_{t(1)} &= \hat{\mathbf{g}}_t^y \mathbf{v}_0, \\
\mathbf{V}_{t(1)} &= \mathbf{D}_0 \hat{\mathbf{q}}_{t(1)}^d,
\end{aligned} \tag{6}$$

where  $\mathbf{q}_{t(1)}^d = \mathbf{U}_{t(1)}^d \mathbf{z} + \mathbf{Y}_{t(1)}^d \mathbf{z}$ . Note that the first four equations in (6) require that the number of industries and the number of commodities have to be equal, otherwise, for example,  $\hat{\mathbf{g}}_t^y \mathbf{U}_0^d$  is not defined. Thus, even if the Euro method distinguishes between the products and industries, it does not allow to estimate rectangular SUTs.

The total *industry* outputs and inputs are *not* equal after this first step, i.e.,  $\mathbf{x}'_{out,t(1)} = \mathbf{z}' \mathbf{V}'_{t(1)} \neq \mathbf{z}' (\mathbf{U}_{t(1)}^d + \mathbf{U}_{t(1)}^m) + \mathbf{v}'_{t(1)} = \mathbf{x}'_{inp,t(1)}$ . To make the derived SUT consistent, it is assumed that the domestic and imported input structures of industries and the totals of commodities' final uses from the first step are valid. Given this assumption, the Euro method uses the so-called fixed commodity sales structure model that determines the consistent output and input levels of industries (Eurostat 2008, Model D, p. 351). Thus, in the second step the consistent industry outputs are derived from the equation

$$\mathbf{x}_{t(2)} = (\mathbf{I} - \mathbf{D}_0 \mathbf{B}_{t(1)}^d)^{-1} \mathbf{D}_0 \mathbf{f}_{t(1)}^d, \tag{7}$$

where the domestic input structure of industries is given by  $\mathbf{B}_{t(1)}^d = \mathbf{U}_{t(1)}^d (\hat{\mathbf{x}}_{inp,t(1)})^{-1}$  and the vector of final demand by product is  $\mathbf{f}_{t(1)}^d = \mathbf{Y}_{t(1)}^d \mathbf{z}$ . Using the imported input structure of industries given by  $\mathbf{B}_{t(1)}^m = \mathbf{U}_{t(1)}^m (\hat{\mathbf{x}}_{inp,t(1)})^{-1}$  and (7), one can easily derive the consistent SUTs components of the first iteration as follows

$$\begin{aligned}
\mathbf{U}_{t(2)}^d &= \mathbf{B}_{t(1)}^d \hat{\mathbf{x}}_{t(2)}, \\
\mathbf{U}_{t(2)}^m &= \mathbf{B}_{t(1)}^m \hat{\mathbf{x}}_{t(2)}, \\
\mathbf{Y}_{t(2)}^d &= \mathbf{Y}_{t(1)}^d, \\
\mathbf{Y}_{t(2)}^m &= \mathbf{Y}_{t(1)}^m, \\
\mathbf{v}'_{t(2)} &= \mathbf{x}'_{t(2)} - \mathbf{z}' (\mathbf{U}_{t(2)}^d + \mathbf{U}_{t(2)}^m), \\
\mathbf{V}_{t(2)} &= \mathbf{D}_0 \hat{\mathbf{q}}_{t(2)}^d,
\end{aligned} \tag{8}$$

where  $\mathbf{q}_{t(2)}^d = \mathbf{U}_{t(2)}^d \mathbf{z} + \mathbf{Y}_{t(2)}^d \mathbf{z}$ . This second step ensures consistency of the sectoral outputs and inputs, and commodity supply and demand.

Deviation between the macroeconomic forecasts and the above projection results is defined as  $dev = (\text{forecasted growth rates})/(\text{projected growth rates})$  of  $ind + f + 2$  variables, which are value added of industries, final uses of  $f$  categories, total value added, and total imports. The growth rates used in (6) are then adjusted in an iterative procedure in order to make the difference between the macroeconomic forecasts and the projected growth rates minimal (less than 1% or even much lower). The observed deviations are used to correct the growth rates  $\mathbf{g}_t^v$  and  $\mathbf{g}_t^y$  employed in (6) according to some rule. This should ensure that if the model overestimates (underestimates) the macroeconomic forecasts, the corresponding growth rates are decreased (increased). For such a purpose the so-called adjustment function of type B is used, according to which the correction factor of variable  $k$ ,  $c_k$ , is defined as follows

$$c_k = \begin{cases} 1 + (\Delta_k \times 100)^\epsilon / 100 & \text{if } \Delta_k > 0, \\ 1 - (-\Delta_k \times 100)^\epsilon / 100 & \text{if } \Delta_k < 0, \end{cases} \quad (9)$$

where  $\Delta_k = dev_k - 1$  and  $\epsilon$  is the adjustment elasticity that is usually set at 0.9. Then the first step of the second iteration computes the projected SUTs components as follows (compare to (6))

$$\begin{aligned} \mathbf{U}_{t(1)}^d &= 0.5 \times (\hat{\mathbf{g}}_t^d \mathbf{U}_0^d + \mathbf{U}_0^d \hat{\mathbf{g}}_t^d), \\ \mathbf{U}_{t(1)}^m &= 0.5 \times (\hat{\mathbf{g}}_t^m \mathbf{U}_0^m + \mathbf{U}_0^m \hat{\mathbf{g}}_t^d), \\ \mathbf{Y}_{t(1)}^d &= 0.5 \times (\hat{\mathbf{g}}_t^d \mathbf{Y}_0^d + \mathbf{Y}_0^d \hat{\mathbf{g}}_t^y), \\ \mathbf{Y}_{t(1)}^m &= 0.5 \times (\hat{\mathbf{g}}_t^m \mathbf{Y}_0^m + \mathbf{Y}_0^m \hat{\mathbf{g}}_t^y), \\ \mathbf{v}_{t(1)} &= \hat{\mathbf{g}}_t^d \mathbf{v}_0, \\ \mathbf{V}_{t(1)} &= \mathbf{D}_0 \hat{\mathbf{q}}_{t(1)}^d, \end{aligned} \quad (10)$$

where the row and column growth rates *adjustment multipliers* are  $\hat{\mathbf{g}}_t^d = \hat{\mathbf{c}}^v \hat{\mathbf{g}}_t^v$  ( $c_i^v$  is the correction factor corresponding to the value added of industry  $i$ ),  $\hat{\mathbf{g}}_t^m = c_m \hat{\mathbf{g}}_t^v$  ( $c_m$  is the correction factor corresponding to total imports), and  $\hat{\mathbf{g}}_t^y = \hat{\mathbf{c}}^y \hat{\mathbf{g}}_t^y$  ( $c_f^y$  is the correction factor of the final demand category  $f$ ).

As was the case with step 1 of the first iteration, the results in (10) do not ensure the equality of industry outputs and inputs. The consistent outputs and inputs of industries are again found using the fixed commodity sales structure model given in (7), which are then used to derive the consistent SUTs of the second iteration in exactly similar way as defined in (8). However, note that now the domestic and imported input structure matrices are derived from the outcomes of the first step of the second iteration. As a result one obtains the new deviation vector of sectoral value added, final demand categories and total imports, which quantifies the difference of the projected growth rates from the macroeconomic forecasts. If

the difference of the growth rates (of sectoral value added, totals of final demand categories, total value added, and total imports) are acceptable (say, less than 1%), the resulting SUT is the final outcome of the Euro projection. Otherwise, the steps of the *second* iteration are repeated until the projected variables resemble (closely or even perfectly) those of the macroeconomic forecasts. It is important to note that each such subsequent iteration begins with computing new correction factors, which are then multiplied by the row and column growth rates adjustment multipliers from the *previous* iteration in order to obtain the corresponding new row and column growth rates adjustment multipliers. For example, the row growth rates adjustment multipliers of domestic uses in iteration  $iter + 1$  are the diagonal elements of the matrix  $\hat{\mathbf{g}}_{t,iter+1}^d = \hat{\mathbf{c}}_{iter+1}^v \hat{\mathbf{g}}_{t,iter}^d$  for all iterations  $iter \geq 2$ . The last important, to our view, point concerning Euro method is that its convergence is *not* always guaranteed. This is in fact expectable, since we do not see any reason from theoretical perspective for its convergence. However, there exist a practical way out of this computative instability problem and it is increasing (slightly) the value of the used tolerance level until convergence can be obtained.

## 2.3 Generalized RAS

The well-known RAS method, which balances a *non-negative* matrix to prescribed row and column totals, is widely used in updating input-output (IO) tables.<sup>7</sup> Its generalization of estimating a matrix that contains both positive and negative entries was first proposed by Günlük-Şenesen and Bates (1988), and later re-discovered by Junius and Oosterhaven (2003) who labeled it as a *generalized* RAS (GRAS) approach. In fact, Günlük-Şenesen and Bates (1988) named the method as a *corrected* RAS method, whose work for some reason was largely unknown, at least, in the recent literature on updating IO tables.<sup>8</sup> The GRAS method presented in Junius and Oosterhaven (2003) was revised by Lenzen et al. (2007) who corrected its biased objective function. However, this correction does not apply to the work of Günlük-Şenesen and Bates (1988) because they set the underlying minimization problem such that it accounts for the bias in the RAS objective function (namely, the row and column constraints with their respective multipliers are taken under the logarithm). We will briefly present the method, while for details the reader is referred to the original sources.

---

<sup>7</sup>The details of the RAS procedure is thoroughly discussed in Miller and Blair (2009, Chapter 7), which also gives an extensive list of references on the topic.

<sup>8</sup>It seems fair to use the original name of the method as a corrected RAS (CRAS, C for corrected) in order to give credit to Günlük-Şenesen and Bates (1988). However, now there is a CRAS (cell-corrected RAS) approach due to Mínguez et al. (2009), hence we follow the current literature and use the term GRAS.

Let  $\mathbf{A}$  be the  $m \times n$  matrix, not necessarily non-negative, with row and column sums of, respectively,  $\mathbf{u}_0 = \mathbf{A}\mathbf{z}$  and  $\mathbf{v}_0 = \mathbf{A}'\mathbf{z}$ . Given ‘new’ row and column sums  $\mathbf{u}$  and  $\mathbf{v}$ , the problem is finding a ‘new’ matrix  $\mathbf{X}$  that is as close as possible to the original matrix  $\mathbf{A}$  and satisfies  $\mathbf{X}\mathbf{z} = \mathbf{u}$  and  $\mathbf{X}'\mathbf{z} = \mathbf{v}$ . Define the ratio of the ‘new’ to the ‘old’ entries by  $z_{ij} \equiv \frac{x_{ij}}{a_{ij}}$  whenever  $a_{ij} \neq 0$ . For  $a_{ij} = 0$ , set  $z_{ij} = 1$ . The appropriate objective function that evaluates the distance between  $\mathbf{A}$  and  $\mathbf{X}$  in the GRAS method is given in Huang et al. (2008).<sup>9</sup> So the problem is

$$\begin{aligned} \min_{z_{ij}} f(\mathbf{Z}) &= \sum_i \sum_j |a_{ij}| (z_{ij} \ln(z_{ij}/e) + 1) \\ \text{such that} & \\ \sum_j a_{ij} z_{ij} &= u_i \quad \text{for all } i = 1, \dots, m, \\ \sum_i a_{ij} z_{ij} &= v_j \quad \text{for all } j = 1, \dots, n, \end{aligned} \tag{11}$$

where  $|a_{ij}|$  is the absolute value of  $a_{ij}$  and  $e$  is the base of natural logarithm. The associated Lagrangean is

$$\begin{aligned} \mathcal{L}(\mathbf{Z}, \boldsymbol{\lambda}, \boldsymbol{\tau}) &= \sum_{(i,j) \in P} a_{ij} (z_{ij} \ln(z_{ij}/e) + 1) - \sum_{(i,j) \in N} a_{ij} (z_{ij} \ln(z_{ij}/e) + 1) \\ &+ \sum_i \lambda_i \left( u_i - \sum_j a_{ij} z_{ij} \right) + \sum_j \tau_j \left( v_j - \sum_i a_{ij} z_{ij} \right), \end{aligned}$$

where  $P$  (resp.  $N$ ) is the set of indices  $(i, j)$  for which  $a_{ij} \geq 0$  (resp.  $a_{ij} < 0$ ), and  $\lambda_i$  and  $\tau_j$  are the Lagrange multipliers. The first-order conditions can be easily derived as

$$\begin{aligned} z_{ij} &= r_i s_j \quad \text{for } a_{ij} \geq 0, \\ z_{ij} &= r_i^{-1} s_j^{-1} \quad \text{for } a_{ij} < 0, \end{aligned}$$

where  $r_i = e^{\lambda_i}$  and  $s_j = e^{\tau_j}$ , thus it always holds that  $z_{ij} > 0$ . Note that when  $a_{ij} = 0$ , then it  $x_{ij} = 0$  as well since then  $z_{ij} = 1$ . From the above optimal conditions it follows that “... the procedure RAS is appropriate for positive elements, but needs to be replaced by  $(1/R)\mathbf{A}(1/S)$  for negative elements” (Günlük-Şenesen and Bates

---

<sup>9</sup>Introducing this function, Huang et al. (2008) call the method as the “Improved GRAS”. We should, however, note that the result of this optimization gives exactly the same outcome as in Lenzen et al. (2007), who use  $f(\mathbf{Z}) = \sum_{i,j} |a_{ij}| z_{ij} \ln(z_{ij}/e)$  instead. The only ‘problem’ with the last function is that for  $z_{ij} = 1$ , i.e., when the initial matrix trivially satisfies the prescribed row and column sums, the function is negative,  $f(\mathbf{Z}) = \sum_{i,j} |a_{ij}| \ln(1/e) = \sum_{i,j} |a_{ij}| (0 - 1) < 0$ . Its modified version gives the function value equal to zero as it should be. However, this adjustment does not play any role in determining the optimal  $\mathbf{Z}$ .

1988, p. 476). Employing the first-order conditions with the row and column totals constraints, one can easily see that the diagonal matrices  $\hat{\mathbf{r}}$  and  $\hat{\mathbf{s}}$  solve the following system of nonlinear equations

$$(\hat{\mathbf{r}}\mathbf{P}\hat{\mathbf{s}} - \hat{\mathbf{r}}^{-1}\mathbf{N}\hat{\mathbf{s}}^{-1})\mathbf{z} = \mathbf{u}, \quad (12)$$

$$\mathbf{z}'(\hat{\mathbf{r}}\mathbf{P}\hat{\mathbf{s}} - \hat{\mathbf{r}}^{-1}\mathbf{N}\hat{\mathbf{s}}^{-1}) = \mathbf{v}', \quad (13)$$

where the matrix  $\mathbf{P}$  (resp.  $\mathbf{N}$ ) contains nonnegative (resp. the absolute value of negative) elements of the original matrix  $\mathbf{A}$ , that is,  $\mathbf{A} = \mathbf{P} - \mathbf{N}$ . The solution of the system (12)-(13) is derived *iteratively*, using a formula for the positive root of the second-order equation (see for details Junius and Oosterhaven 2003). Furthermore, since  $z_{ij} > 0$ , the signs of all the elements in the original matrix  $\mathbf{A}$  will be preserved in the projected matrix  $\mathbf{X}$ . Notice that when  $\mathbf{A}$  is non-negative, then  $\mathbf{P} = \mathbf{A}$  and thus  $\mathbf{X} = \hat{\mathbf{r}}\mathbf{A}\hat{\mathbf{s}}$ , which is exactly how the traditional RAS method works.

As in the original RAS approach, the row and column multipliers have the following economic meaning. Multipliers in  $\hat{\mathbf{r}}$  are interpreted as the “substitution factors” because they measure the degree to which the input has replaced or has been replaced over time by other inputs. The column multipliers in  $\hat{\mathbf{s}}$ , on the other hand, are known as the “fabrication factors” since they measure the extent to which the output in each sector has absorbed over time more or less the intermediate inputs vs. primary inputs. In our empirical application of GRAS for updating SUTs, we use the iterative algorithm proposed by Lenzen et al. (2007, p. 465).

## 2.4 Improved normalized squared differences

Friedlander (1961) proposed to use the normalized squared differences as an objective function in order to estimate a matrix with given row and column totals, which has been suggested by Lecomber (1975) as a useful technique for updating IO tables. The objective function is

$$f = \sum_i \sum_j \frac{(x_{ij} - a_{ij})^2}{a_{ij}}, \quad (14)$$

where  $a_{ij} > 0$ . Note that smaller  $a_{ij}$ s have higher contribution to the function (14), hence minimization of this function implies greater penalty on the change of small elements. This means that an updating procedure with the objective (14) results in the matrix whose changes are more concentrated in the larger entries. Huang et al. (2008) rewrite (14) with respect to  $z_{ij} = \frac{x_{ij}}{a_{ij}}$  that also allows for non-positive entries in the original matrix  $\mathbf{A}$  as

$$f(\mathbf{Z}) = \sum_i \sum_j |a_{ij}|(z_{ij} - 1)^2, \quad (15)$$

which they call as an improved normalized squared differences (INSD).

The INSD function is used together with the penalty function of the form  $\frac{M}{2} \sum_i \sum_j |a_{ij}| [\min(0, z_{ij})]^2$ , where  $M > 0$  is the penalty number that prevents  $z_{ij}$  from being significantly negative. If there is  $z_{ij} < 0$ , then for  $M \rightarrow \infty$  this element becomes virtually zero, hence practically the inequality  $z_{ij} \geq 0$  is satisfied. The last inequality guarantees preserving the sign of each entry in  $\mathbf{X}$  as in  $\mathbf{A}$ . Using the constraints as in the GRAS problem (11), the associated Lagrangian is

$$\begin{aligned} \mathcal{L}(\mathbf{Z}, \boldsymbol{\lambda}, \boldsymbol{\tau}) = & \frac{1}{2} \sum_i \sum_j |a_{ij}| ((z_{ij} - 1)^2 + M[\min(0, z_{ij})]^2) \\ & + \sum_i \lambda_i \left( u_i - \sum_j a_{ij} z_{ij} \right) + \sum_j \tau_j \left( v_j - \sum_i a_{ij} z_{ij} \right). \end{aligned}$$

The optimality condition  $\partial \mathcal{L} / \partial z_{ij} = 0$  yields

$$z_{ij} + M \min(0, z_{ij}) = 1 + \frac{a_{ij}}{|a_{ij}|} (\lambda_i + \tau_j),$$

thus

$$z_{ij} = \begin{cases} 1 & \text{if } a_{ij} = 0, \\ 1 + \frac{a_{ij}}{|a_{ij}|} (\lambda_i + \tau_j) & \text{if } 1 + \frac{a_{ij}}{|a_{ij}|} (\lambda_i + \tau_j) \geq 0, \\ \frac{1}{1+M} \left[ 1 + \frac{a_{ij}}{|a_{ij}|} (\lambda_i + \tau_j) \right] & \text{otherwise.} \end{cases} \quad (16)$$

Next, using the row and column constraints of  $\mathbf{X}$  it can be easily shown that the Lagrange multipliers are equal to (see for details Huang et al. 2008, pp. 116-118)

$$\lambda_i = \frac{1}{\sum_j |a_{ij}|} \left[ u_i - \sum_j a_{ij} + \sum_j \{ M a_{ij} \min(0, z_{ij}) - \tau_j |a_{ij}| \} \right], \quad (17)$$

$$\tau_j = \frac{1}{\sum_i |a_{ij}|} \left[ v_j - \sum_i a_{ij} + \sum_i \{ M a_{ij} \min(0, z_{ij}) - \lambda_i |a_{ij}| \} \right]. \quad (18)$$

Equations (16)-(18) are used in an iterative procedure to obtain the final solution  $\mathbf{Z}$ . We initialize  $\mathbf{Z}_0 = \mathbf{U}$  and  $\boldsymbol{\lambda}_0 = \boldsymbol{\tau}_0 = \mathbf{0}$ , where  $\mathbf{U}$  is the  $m \times n$  matrix of ones and  $\mathbf{0}$  is the null vector. Then at step  $t = 1, 2, \dots, K$  using (17)-(18) we obtain  $\boldsymbol{\lambda}_t$  that is used to derive  $\boldsymbol{\tau}_t$ , which are then employed to obtain  $\mathbf{Z}_t$  from (16). This process continues until for some finite  $K$  (if the solution exists)  $\boldsymbol{\lambda}_K - \boldsymbol{\lambda}_{K-1} < \varepsilon \mathbf{1}$  and  $\boldsymbol{\tau}_K - \boldsymbol{\tau}_{K-1} < \varepsilon \mathbf{1}$  for sufficiently small  $\varepsilon > 0$ . The final solution  $\mathbf{Z}$  is derived using the last step multipliers  $\boldsymbol{\lambda}_K$  and  $\boldsymbol{\tau}_K$ , from which the matrix  $\mathbf{X}$  is obtained via  $x_{ij} = z_{ij} a_{ij}$ .

As observed by Huang et al. (2008), the (improved) GRAS function is an approximate expression of the INSD function: if we take the first-order Taylor series



of  $\ln(z_{ij})$  around  $z_{ij} = 1$ , the GRAS function boils down to the INSD objective, i.e.,  $|a_{ij}|(z_{ij} \ln(z_{ij}/e) + 1) \approx |a_{ij}|(0 + z_{ij}(z_{ij} - 2) + 1) = |a_{ij}|(z_{ij} - 1)^2$ .<sup>10</sup> Because of this closeness and the fact that in GRAS it always holds that  $z_{ij} > 0$ , the “INSD is more likely to preserve the signs of entries than other non-biproportional methods, similar to [GRAS]” (Huang et al. 2008, p. 121).

## 2.5 Improved squared differences

Another often used objective for updating IO matrices is the function of squared differences of Almon (1968)

$$f = \sum_i \sum_j (x_{ij} - a_{ij})^2. \quad (19)$$

Huang et al. (2008) rewrite (19) in terms of  $z_{ij}$  to derive the improved squared differences (ISD) function

$$f(\mathbf{Z}) = \sum_i \sum_j a_{ij}^2 (z_{ij} - 1)^2, \quad (20)$$

which has the same form as the sign preserving squared differences in Jackson and Murray (2004). Again to ‘insure’ non-negativity of  $z_{ij}$ , the penalty function of the form  $\frac{M}{2} \sum_i \sum_j a_{ij}^2 [\min(0, z_{ij})]^2$  is used together with the ISD function in (20). This function subject to the usual prescribed row and column constraints in (11) yields the Lagrangian

$$\begin{aligned} \mathcal{L}(\mathbf{Z}, \boldsymbol{\lambda}, \boldsymbol{\tau}) = & \frac{1}{2} \sum_i \sum_j a_{ij}^2 ((z_{ij} - 1)^2 + M[\min(0, z_{ij})]^2) \\ & + \sum_i \lambda_i \left( u_i - \sum_j a_{ij} z_{ij} \right) + \sum_j \tau_j \left( v_j - \sum_i a_{ij} z_{ij} \right). \end{aligned}$$

The optimality condition  $\partial \mathcal{L} / \partial z_{ij} = 0$  gives

$$z_{ij} + M \min(0, z_{ij}) = 1 + \frac{\lambda_i + \tau_j}{a_{ij}},$$

thus

$$z_{ij} = \begin{cases} 1 & \text{if } a_{ij} = 0, \\ 1 + \frac{\lambda_i + \tau_j}{a_{ij}} & \text{if } 1 + \frac{\lambda_i + \tau_j}{a_{ij}} \geq 0, \\ \frac{1}{1+M} \left[ 1 + \frac{\lambda_i + \tau_j}{a_{ij}} \right] & \text{otherwise.} \end{cases} \quad (21)$$

---

<sup>10</sup>An equivalent closeness between the normalized squared differences and RAS outcomes have been already mentioned in Harrigan and Buchanan (1984, p. 341).

Using the constraints  $\sum_j x_{ij} = u_i$  and  $\sum_i x_{ij} = v_j$  together with the above optimality condition (when  $a_{ij} \neq 0$ ) yields

$$\lambda_i = \frac{1}{\sum_j \delta(a_{ij})} \left[ u_i - \sum_j a_{ij} + \sum_j \{ M a_{ij} \min(0, z_{ij}) - \tau_j \delta(a_{ij}) \} \right], \quad (22)$$

$$\tau_j = \frac{1}{\sum_i \delta(a_{ij})} \left[ v_j - \sum_i a_{ij} + \sum_i \{ M a_{ij} \min(0, z_{ij}) - \lambda_i \delta(a_{ij}) \} \right], \quad (23)$$

where  $\delta(a_{ij}) = 1$  if  $a_{ij} \neq 0$ , and  $\delta(a_{ij}) = 0$  otherwise. The iterative algorithm is similar to the one when the objective of the constrained minimization problem is the INSD and is described in the previous section.

## 2.6 Improved weighted squared differences

The errors in the squared differences objective (19) are equally weighted despite the size of the original entries of  $\mathbf{A}$ . Some analysts would argue that large coefficients are the outcome of more permanent and stronger linkages, and thus are more stable than small coefficients (see e.g., Hinojosa 1978). The squared differences version of the weighted absolute deviation measure in Lahr (2001) can be written as (see also Pavia et al. 2009)

$$f = \sum_i \sum_j a_{ij} (x_{ij} - a_{ij})^2, \quad (24)$$

where  $a_{ij} > 0$ . Thus when the squared differences are weighted by the corresponding original coefficients, in contrast to the normalized squared differences (14), changes in the larger entries are penalized more in the minimization of (24). Hence, it is expected that changes in the updated matrix are more concentrated in the smaller entries.

To allow also for negative entries, rewrite (24) with respect to  $z_{ij} = \frac{x_{ij}}{a_{ij}}$  as

$$f(\mathbf{Z}) = \sum_i \sum_j |a_{ij}^3| (z_{ij} - 1)^2, \quad (25)$$

which can be called (in conjunction with ISD and INSD) as the improved weighted squared differences (IWSD). Again for the sign preservation, we use now the penalty function  $\frac{M}{2} \sum_i \sum_j |a_{ij}^3| [\min(0, z_{ij})]^2$ , so that the associated Lagrangian is

$$\begin{aligned} \mathcal{L}(\mathbf{Z}, \boldsymbol{\lambda}, \boldsymbol{\tau}) &= \frac{1}{2} \sum_i \sum_j |a_{ij}^3| \left( (z_{ij} - 1)^2 + M [\min(0, z_{ij})]^2 \right) \\ &\quad + \sum_i \lambda_i \left( u_i - \sum_j a_{ij} z_{ij} \right) + \sum_j \tau_j \left( v_j - \sum_i a_{ij} z_{ij} \right). \end{aligned}$$

The optimality condition  $\partial\mathcal{L}/\partial z_{ij} = 0$  is  $z_{ij} + M \min(0, z_{ij}) = 1 + \frac{a_{ij}(\lambda_i + \tau_j)}{|a_{ij}^3|}$ , thus

$$z_{ij} = \begin{cases} 1 & \text{if } a_{ij} = 0, \\ 1 + \frac{a_{ij}(\lambda_i + \tau_j)}{|a_{ij}^3|} & \text{if } 1 + \frac{a_{ij}(\lambda_i + \tau_j)}{|a_{ij}^3|} \geq 0, \\ \frac{1}{1+M} \left[ 1 + \frac{a_{ij}(\lambda_i + \tau_j)}{|a_{ij}^3|} \right] & \text{otherwise.} \end{cases} \quad (26)$$

Employing the row and column constraints of  $\mathbf{X}$  it can be easily shown that the Lagrange multipliers are

$$\lambda_i = \frac{1}{\sum_j (|a_{ij}|^{-1})} \left[ u_i - \sum_j a_{ij} + \sum_j \left\{ M a_{ij} \min(0, z_{ij}) - \frac{\tau_j}{|a_{ij}|} \right\} \right], \quad (27)$$

$$\tau_j = \frac{1}{\sum_i (|a_{ij}|^{-1})} \left[ v_j - \sum_i a_{ij} + \sum_i \left\{ M a_{ij} \min(0, z_{ij}) - \frac{\lambda_i}{|a_{ij}|} \right\} \right], \quad (28)$$

where we define  $\frac{1}{|a_{ij}|} = 1$  for  $a_{ij} = 0$ . Equations (26)-(28) are used in an iterative algorithm (similar to the INSD algorithm described in Section 2.4) to derive the final solution  $\mathbf{X}$ .

## 2.7 Harthoorn and van Dalen's method

A more general form of the least squares 'distance' approach was proposed by Harthoorn and van Dalen (1987). They set the matrix updating problem by introducing factors  $f_{ij}$  for each entry of the original  $m \times n$  matrix  $\mathbf{A}$  which determine the updated matrix  $\mathbf{X}$  as follows:

$$\begin{aligned} \min_{f_{ij}} f(\mathbf{F}) &= \sum_i \sum_j (f_{ij} a_{ij} - a_{ij})^2 / g_{ij} \\ \text{such that} & \\ \sum_j f_{ij} a_{ij} &= u_i \quad \text{for all } i = 1, \dots, m, \\ \sum_i f_{ij} a_{ij} &= v_j \quad \text{for all } j = 1, \dots, n, \end{aligned} \quad (29)$$

where  $g_{ij}$  is the known relative confidence of element  $a_{ij}$ . Since this problem does not guarantee sign-preservation (some  $f_{ij}$  can be negative), similar to the previous problems, we introduce the penalty function  $\frac{M}{2} \sum_i \sum_j \frac{a_{ij}^2}{g_{ij}} [\min(0, f_{ij})]^2$ . Thus, the Lagrangian becomes

$$\begin{aligned}\mathcal{L}(\mathbf{F}, \boldsymbol{\lambda}, \boldsymbol{\tau}) = & \frac{1}{2} \sum_i \sum_j \frac{a_{ij}^2}{g_{ij}} ((f_{ij} - 1)^2 + M[\min(0, f_{ij})]^2) \\ & + \sum_i \lambda_i \left( u_i - \sum_j f_{ij} a_{ij} \right) + \sum_j \tau_j \left( v_j - \sum_i f_{ij} a_{ij} \right).\end{aligned}$$

The optimality condition  $\partial \mathcal{L} / \partial f_{ij} = 0$  gives

$$f_{ij} + M \min(0, f_{ij}) = 1 + \frac{g_{ij}(\lambda_i + \tau_j)}{a_{ij}},$$

thus

$$f_{ij} = \begin{cases} 1 & \text{if } a_{ij} = 0, \\ 1 + \frac{g_{ij}(\lambda_i + \tau_j)}{a_{ij}} & \text{if } 1 + \frac{g_{ij}(\lambda_i + \tau_j)}{a_{ij}} \geq 0, \\ \frac{1}{1+M} \left[ 1 + \frac{g_{ij}(\lambda_i + \tau_j)}{a_{ij}} \right] & \text{otherwise.} \end{cases} \quad (30)$$

Note that in contrast to the least squares objectives, we could have defined  $f_{ij} = 0$  for  $a_{ij} = 0$  because from the outset it was not defined to be the ratio of the new and old entries like  $z_{ij}$ .

Using the constraints from (29) together with the above optimality condition gives

$$\lambda_i = \frac{1}{\sum_j g_{ij}} \left[ u_i - \sum_j a_{ij} + \sum_j \{ M a_{ij} \min(0, z_{ij}) - g_{ij} \tau_j \} \right], \quad (31)$$

$$\tau_j = \frac{1}{\sum_i g_{ij}} \left[ v_j - \sum_i a_{ij} + \sum_i \{ M a_{ij} \min(0, z_{ij}) - g_{ij} \lambda_i \} \right]. \quad (32)$$

If the weights are chosen such that  $g_{ij} = 1$  for all  $i$  and all  $j$  when  $a_{ij} \neq 0$ , and zero otherwise, then it is easy to see that the ISD becomes a particular case of the Harthoorn and van Dalen's (HvD) method (compare (30)-(32) to (21)-(23)). The authors call 'additive adjustment' the case when  $g_{ij} = 1$  for all  $i$  and all  $j$  (also for  $a_{ij} = 0$ ). Harthoorn and van Dalen (1987) suggest two other specifications of the confidence intervals  $g_{ij}$ : (i) "[i]f the matrix to be adjusted in is based on a relatively small unbiased sample, it often turns out that the squares of the elements,  $a_{ij}^2$ , should be taken", and (ii) "[i]t is not unusual for elements with larger values to be estimated more accurately in the compilation of input-output tables. This varying accuracy could be expressed by choosing the elements themselves as weights. As this would lead to an adjustment in the wrong direction in the case of negative elements, the absolute values  $|a_{ij}|$  should be taken" (pp. 11-12). Note that with  $g_{ij} = |a_{ij}|$ , the objective in (29) boils down to the INSD function (15), hence INSD is a particular

case of the HvD method. Equations (30)-(32) are used in an iterative algorithm, similar to the INSD algorithm described in Section 2.4, to derive at the final solution of  $\mathbf{X}$ .

## 2.8 Kuroda's method

Kuroda (1988) proposed the following objective function in updating IO matrices

$$f = \frac{1}{2} \sum_i \sum_j \left[ \left( \frac{x_{ij}}{u_i} - r_{ij} \right)^2 w_{ij} + \left( \frac{x_{ij}}{v_j} - c_{ij} \right)^2 v_{ij} \right], \quad (33)$$

where the shares of each element of the original matrix in the row and column sums are  $r_{ij} = \frac{a_{ij}}{u_i^0}$  and  $c_{ij} = \frac{a_{ij}}{v_j^0}$  (superscript 0 refers to the base year, i.e.,  $u_i^0 = \sum_j a_{ij}$  and  $v_j^0 = \sum_i a_{ij}$ ), and  $w_{ij}$  and  $v_{ij}$  are arbitrary sets of positive weights. Further, from (33) follows that  $u_i^0$ ,  $u_i$ ,  $v_j^0$  and  $v_j$  are assumed to be nonzero for all  $i$  and all  $j$ . We rewrite (33) in terms of  $z_{ij} = \frac{x_{ij}}{a_{ij}}$  as follows

$$f(\mathbf{Z}) = \frac{1}{2} \sum_i \sum_j a_{ij}^2 \left[ \left( \frac{z_{ij}}{u_i} - \frac{1}{u_i^0} \right)^2 w_{ij} + \left( \frac{z_{ij}}{v_j} - \frac{1}{v_j^0} \right)^2 v_{ij} \right]. \quad (34)$$

Note that the ISD function in (20) is a particular case of (34) with  $w_{ij} = v_{ij} = u_i = u_i^0 = v_i = v_i^0 = 1$  for all  $i$  and all  $j$ . For simplicity we denote

$$s_{ij} \equiv \frac{w_{ij}}{u_i^2} + \frac{v_{ij}}{v_j^2} > 0,$$

and note that with Kuroda's method it must be the case that the row and column sums of the matrices  $\mathbf{A}$  and  $\mathbf{X}$  are nonzero.

Similar to Huang et al. (2008), we use the penalty function together with (34) in order to prevent  $z_{ij}$  from being significantly negative, which in current setting has the form  $\frac{M}{2} \sum_i \sum_j a_{ij}^2 s_{ij} [\min(0, z_{ij})]^2$ . Considering also the row and column constraints in (11), the associated Lagrangian becomes

$$\begin{aligned} \mathcal{L}(\mathbf{Z}, \boldsymbol{\lambda}, \boldsymbol{\tau}) = & \frac{1}{2} \sum_i \sum_j a_{ij}^2 \left[ \left( \frac{z_{ij}}{u_i} - \frac{1}{u_i^0} \right)^2 w_{ij} + \left( \frac{z_{ij}}{v_j} - \frac{1}{v_j^0} \right)^2 v_{ij} + M s_{ij} [\min(0, z_{ij})]^2 \right] \\ & + \sum_i \lambda_i \left( u_i - \sum_j a_{ij} z_{ij} \right) + \sum_j \tau_j \left( v_j - \sum_i a_{ij} z_{ij} \right). \end{aligned}$$

The optimality condition  $\partial \mathcal{L} / \partial z_{ij} = 0$  gives

$$z_{ij} + M \min(0, z_{ij}) = \frac{1}{s_{ij}} \left( s_{ij}^0 + \frac{\lambda_i + \tau_j}{a_{ij}} \right),$$

where  $s_{ij}^0 \equiv \frac{w_{ij}}{u_i^0 u_i} + \frac{v_{ij}}{v_j^0 v_j}$ . Thus,

$$z_{ij} = \begin{cases} 1 & \text{if } a_{ij} = 0, \\ \frac{1}{s_{ij}} \left( s_{ij}^0 + \frac{\lambda_i + \tau_j}{a_{ij}} \right) & \text{if } \frac{1}{s_{ij}} \left( s_{ij}^0 + \frac{\lambda_i + \tau_j}{a_{ij}} \right) \geq 0, \\ \frac{1}{(1+M)s_{ij}} \left( s_{ij}^0 + \frac{\lambda_i + \tau_j}{a_{ij}} \right) & \text{otherwise.} \end{cases} \quad (35)$$

Using the constraints  $\sum_j a_{ij} z_{ij} = u_i$  and  $\sum_i a_{ij} x_{ij} = v_j$  together with the above first-order conditions yields

$$\lambda_i = \frac{1}{\sum_j s_{ij}^{-1}} \left[ u_i - \sum_j a_{ij} \frac{s_{ij}^0}{s_{ij}} + \sum_j \left\{ M a_{ij} \min(0, z_{ij}) - \frac{\tau_j}{s_{ij}} \right\} \right], \quad (36)$$

$$\tau_j = \frac{1}{\sum_i s_{ij}^{-1}} \left[ v_j - \sum_i a_{ij} \frac{s_{ij}^0}{s_{ij}} + \sum_i \left\{ M a_{ij} \min(0, z_{ij}) - \frac{\lambda_i}{s_{ij}} \right\} \right]. \quad (37)$$

Equations (35)-(37) are used in an iterative process to obtain the final solution  $\mathbf{Z}$  and  $\mathbf{X}$ . The iterative algorithm is similar to those when the ISND and ISD functions are the objectives in the constrained minimization problem (11) (see Section 2.4).

Note that for the computation of (35)-(37), one needs to choose the weights  $w_{ij}$  and  $v_{ij}$  (that in turn determine the values of  $s_{ij}$  and  $s_{ij}^0$ ). One approach is to weight all errors equally, i.e., set  $w_{ij} = v_{ij} = 1$  for all  $i$  and all  $j$ . Kuroda (1988) proposes “the equal percentage change” weighting scheme of the form

$$w_{ij} = \frac{1}{r_{ij}^2}, \quad v_{ij} = \frac{1}{c_{ij}^2}. \quad (38)$$

Wilcoxon (1989), on the other hand, for the simplification of obtained expressions also considers the following weights

$$w_{ij} = \frac{u_i^2}{2}, \quad v_{ij} = \frac{v_j^2}{2}, \quad (39)$$

which result in  $s_{ij} = 1$  for all  $i$  and all  $j$ .

If one uses (38), the objective function (33) becomes

$$f = \frac{1}{2} \sum_i \sum_j \left[ \left( \frac{x_{ij}/u_i}{r_{ij}} - 1 \right)^2 + \left( \frac{x_{ij}/v_j}{c_{ij}} - 1 \right)^2 \right], \quad (40)$$

which simply represents the percentage changes in the coefficients of the new and original matrices. Notice that Kuroda’s weights in (38) imply that the matrix  $\mathbf{A}$  to be updated should not contain zero elements, otherwise the weights are not defined. In our setting with Kuroda’s weights we do allow both for zero and negative elements

**Table 1:** Short description of the methods for updating SUTs

Method	Source	Data requirements for the projected year(s)
1 EUKLEMS method	Timmer et al. (2005) and further Timmer’s guidelines	Column totals
2 Euro method	Beutel (2002), Eurostat (2008)	Growth rates of: (i) sectoral value added, (ii) totals of final demand categories, and (iii) total imports
3 GRAS (Generalized RAS)	Günlük-Şenesen and Bates (1988), Junius and Oosterhaven (2003)	Row and column totals
4 INSD (Improved normalized squared differences)	Friedlander (1961), Huang et al. (2008)	Row and column totals
5 ISD (Improved squared differences)	Almon (1968), Jackson and Murray (2004), Huang et al. (2008)	Row and column totals
6 IWSD (Improved weighted squared differences)	Pavia et al. (2009) – slightly revised here to allow for negative entries	Row and column totals
7 Harthoorn and van Dalen’s method (INSD and two other types depending on the weights choice)	Harthoorn and van Dalen (1987) – slightly revised here to allow for negative entries	Row and column totals
8 Kuroda’s method (three types depending on the weights choice)	Kuroda (1988) – slightly revised here to allow for negative entries; Wilcoxon (1989)	Row and column totals

*Note:* The row and column totals of the entire projected matrix in the text were denoted, respectively, by the column vectors  $\mathbf{u}$  and  $\mathbf{v}$ . All the methods additionally require availability of the full benchmark matrix  $\mathbf{A}$  of the base year.

as well, and for all  $(k, h)$  that have  $a_{kh} = 0$ , we redefine  $w_{kh} = v_{kh} = s_{kh}^{-1} = 0$ .

The short summary of SUTs and IO updating methods discussed in this work is given in Table 1.

### 3 Empirical evaluation

Each updating method discussed in the previous section produces different set of estimates, thus it is desirable to assess their relative performance. We do *not* use any of the objective functions from Section 2 as the matrix-comparison statistics of the closeness of the estimates to actual matrices, which is, in fact, often practiced in the related literature. We find it a rather strange way of comparison, because it is obvious that if, for example, the GRAS function is chosen as a measure of closeness, then most probably the GRAS estimate will produce the ‘best’ outcome. Instead the closeness statistics have to be ‘neutral’ with respect to all the methods used

and should assess the methods on equal ground. It is evident that choosing a single statistics as a measure of closeness would be misleading either, since it very well may happen that a method can score well on one dimension but very bad on the other. Thus, several measures are used in order to get more insight into the characteristics of each method. We use the following criteria:

1. Mean absolute percentage error (Butterfield and Mules 1980):

$$MAPE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \frac{|x_{ij} - x_{ij}^{true}|}{|x_{ij}^{true}|} \times 100,$$

where  $x_{ij}^{true}$  is the true element, while  $x_{ij}$  is its estimate. Thus,  $MAPE$  shows the *average* percentage by which each estimated element is larger or smaller than its true value. For  $x_{ij}^{true} = 0$  we set the corresponding difference to zero (as zero is preserved in all the estimates). Note that we take the denominator in absolute value as well so that it does not allow to reduce the actual error when  $x_{ij}^{true} < 0$ .

2. Weighted absolute percentage error (Mínguez et al. 2009):

$$WAPE = \sum_{i=1}^m \sum_{j=1}^n \left( \frac{|x_{ij}^{true}|}{\sum_k \sum_l x_{kl}^{true}} \right) \frac{|x_{ij} - x_{ij}^{true}|}{|x_{ij}^{true}|} \times 100,$$

which weights each percentage deviation of  $x_{ij}$  from  $x_{ij}^{true}$  by the relative size of the corresponding true element in the overall sum of the actual elements.

3. Standardized weighted absolute difference (Lahr 2001):

$$SWAD = \frac{\sum_{i=1}^m \sum_{j=1}^n |x_{ij}^{true}| \times |x_{ij} - x_{ij}^{true}|}{\sum_k \sum_l (x_{kl}^{true})^2},$$

which is effectively similar to  $WAPE$  with the difference that the absolute deviations are weighted by the size of the true transactions.

4. The *psi statistic* (Kullback 1959, Knudsen and Fotheringham 1986):

$$\hat{\psi} = \frac{1}{\sum_k \sum_l x_{kl}^{true}} \sum_i \sum_j \left[ |x_{ij}^{true}| \times \left| \ln \left( \frac{x_{ij}^{true}}{s_{ij}} \right) \right| + |x_{ij}| \times \left| \ln \left( \frac{x_{ij}}{s_{ij}} \right) \right| \right],$$

where  $s_{ij} = (|x_{ij}^{true}| + |x_{ij}|)/2$ . This information-based statistic has a lower limit of zero when  $\mathbf{X}^{true} = \mathbf{X}$ , and upper bound of  $mn \ln 2$  when the non-zero elements of  $\mathbf{X}^{true}$  correspond to the zero elements of  $\mathbf{X}$ , and vice versa. Unlike  $MAPE$ ,  $WAPE$  and  $SWAD$ , the psi statistic is insensitive to the change in the positions of  $x_{ij}^{true}$  and  $x_{ij}$ , and it offers the advantage of considering the case when  $x_{ij}^{true} = 0$  and  $x_{ij} \neq 0$  (next to the reverse situation).<sup>11</sup> Knudsen and Fotheringham (1986) concluded that

<sup>11</sup>When  $x_{ij}^{true} = 0$ , we set the corresponding element of  $MAPE$ ,  $WAPE$  and  $SWAD$  to zero,



the psi statistic is one of the most useful goodness-of-fit measures for comparative purposes because  $\hat{\psi}$  indicator showed a linear relation between its value and the level of error.

5. *RSQ* (or coefficient of determination) – the square of the correlation coefficient between the elements of the actual and predicted matrices of  $\mathbf{X}^{true}$  and  $\mathbf{X}$ , respectively, when at least one of them is different from zero.

6. *N0* - number of zero elements in the estimated matrix  $\mathbf{X}$ , whose corresponding elements are nonzero in the actual matrix  $\mathbf{X}^{true}$ , i.e., the count of  $(i, j)$ s with  $x_{ij} = 0$  for which  $x_{ij}^{true} \neq 0$ . This situation happens for two reasons: first, over time zero transactions can turn into positive links, which represents ‘structural change’ and larger interdependencies in the economy over time. Second, some updating methods nullify certain elements in  $\mathbf{X}$ , which is the case when a change in the sign of the predicted entries is expected, but the corresponding penalty functions instead make them (virtually) zero. Both these effects can be easily found from *N0* as will be discussed below.

In our first empirical application, we use three benchmark Supply and Use tables of the Netherlands for the fiscal years of 1995, 2000, and 2005. There are 59 commodities and 60 industries. For the supply matrices we deleted four industries with zero intermediate totals, and one commodity without any industrial use, hence we used and estimated  $58 \times 56$  Supply tables. Apart from that, for the Use tables, we use four final demand categories: final consumption expenditure by households, consumption by non-profit organizations serving households (NPISH), government expenditures, and gross fixed capital formation. We did not predict commodity exports and imports (in the Supply table), since these data are often available from the trade statistics. Further, deleting one extra row of product without any use, we ended up with  $57 \times 59$  Use tables.<sup>12</sup>

The results of updating Supply and Use matrices are given, respectively, in Table 2 and Table 3.<sup>13</sup> Let consider the upper block of the first table, which gives the results of the updated Supply table for 2000 on the basis of the 1995 Supply table, and the row and column sums from 2000 Supply table. We see that Harthoorn and van Dalen’s method with weights being the squares of the original entries  $a_{ij}^2$  (HvD1) is, on average, 10.66% ‘in error’ according to *MAPE*. On the other hand, the least error according to *WAPE* is caused by RAS updating. The number *N0* for the first two methods (i.e., EUKLEMS and RAS) will always show the count of

---

and when  $x_{ij}^{true} = x_{ij} = 0$ , the corresponding entry of  $\hat{\psi}$  is nullified as well.

<sup>12</sup>In the 2005 Use table there was the third commodity Retail trade services without any use, but we did not take it away, since these services were provided to the households at the amount of 492 and 461 Mln Euros for years 1995 and 2000, respectively. As a result we add a small number (unity) in the corresponding cell in the 2005 Use table to make the programs ‘feasible’.

<sup>13</sup>All the projections are implemented in MATLAB.

**Table 2:** Results of updating Supply tables at basic prices for the Netherlands

	MAPE	R.	WAPE	R.	SWAD	R.	$\hat{\psi}$	R.	RSQ	R.	N0	R.	CmR.
1995 $\Rightarrow$ 2000													
EUKLEMS	12.95	5	4.17	6	0.010	6	0.041	6	0.9992	6	28	1	5
RAS	12.76	4	3.21	1	0.005	2	0.031	1	0.9997	1	28	1	1
HvD1	10.66	1	4.71	7	0.011	7	0.046	7	0.9992	6	31	5	6
HvD/INSD	12.46	3	3.30	2	0.005	1	0.032	2	0.9997	1	28	1	1
HvD3	137.20	8	29.33	9	0.229	9	0.330	9	0.9800	8	85	9	9
ISD	271.85	10	37.92	10	0.201	8	0.346	10	0.9753	10	200	10	10
IWSD	266.89	9	27.14	8	0.236	10	0.262	8	0.9775	9	53	7	8
Kuroda1	12.07	2	3.70	3	0.006	3	0.036	3	0.9995	4	29	4	3
Kuroda2	15.18	7	3.78	5	0.008	5	0.037	5	0.9995	4	62	8	7
Kuroda3	14.99	6	3.72	4	0.007	4	0.036	4	0.9996	3	50	6	4
0.5 $\times$ (1995 + 2005) $\Rightarrow$ 2000													
EUKLEMS	14.17	5	5.23	7	0.016	7	0.051	7	0.9987	7	10	1	6
RAS	14.05	4	4.27	4	0.012	4	0.042	4	0.9994	4	10	1	4
HvD1	12.87	1	4.01	2	0.009	2	0.040	2	0.9995	1	11	3	2
HvD/INSD	13.75	3	4.09	3	0.010	3	0.040	3	0.9995	1	11	3	3
HvD3	39.06	9	10.21	10	0.060	9	0.099	10	0.9935	9	97	9	10
ISD	89.41	10	9.43	8	0.031	8	0.088	8	0.9978	8	182	10	8
IWSD	38.68	8	10.09	9	0.062	10	0.099	9	0.9931	10	55	8	9
Kuroda1	13.14	2	3.71	1	0.008	1	0.037	1	0.9995	1	11	3	1
Kuroda2	21.04	7	4.91	6	0.016	6	0.048	6	0.9990	6	43	7	7
Kuroda3	17.57	6	4.78	5	0.016	5	0.046	5	0.9991	5	38	6	5
1995 $\Rightarrow$ 2005													
EUKLEMS	26.25	5	8.45	7	0.028	6	0.079	7	0.9964	7	73	1	7
RAS	26.14	4	6.25	2	0.014	1	0.059	2	0.9985	1	73	1	2
HvD1	18.98	1	8.03	6	0.034	7	0.075	6	0.9975	6	76	4	5
HvD/INSD	24.56	3	6.48	3	0.021	3	0.061	3	0.9978	3	80	5	3
HvD3	219.22	8	48.45	9	0.396	9	0.461	9	0.9529	8	177	9	9
ISD	432.07	10	60.85	10	0.333	8	0.548	10	0.9277	10	288	10	10
IWSD	341.14	9	44.96	8	0.410	10	0.430	8	0.9442	9	111	7	8
Kuroda1	23.48	2	6.02	1	0.021	2	0.056	1	0.9978	3	73	1	1
Kuroda2	33.51	6	7.40	5	0.023	5	0.069	5	0.9978	3	172	8	6
Kuroda3	35.01	7	7.11	4	0.022	4	0.066	4	0.9980	2	108	6	4
(1995 $\Rightarrow$ 2000) $\Rightarrow$ 2005													
EUKLEMS	26.25	5	8.45	6	0.028	6	0.079	6	0.9964	7	73	1	5
RAS	26.14	4	6.25	1	0.014	1	0.059	1	0.9985	1	73	1	1
HvD1	19.06	1	8.58	7	0.037	7	0.081	7	0.9969	6	76	5	6
HvD/INSD	24.50	3	6.54	3	0.022	3	0.061	3	0.9977	4	73	1	2
HvD3	870.06	9	50.34	9	0.393	9	0.486	8	0.9528	8	131	8	8
ISD	543.08	8	63.06	10	0.335	8	0.613	10	0.9272	9	151	9	10
IWSD	1118.10	10	49.10	8	0.410	10	0.608	9	0.8904	10	88	6	9
Kuroda1	23.59	2	6.50	2	0.024	5	0.061	2	0.9975	5	74	4	3
Kuroda2	48.06	7	7.44	5	0.023	4	0.071	5	0.9978	3	152	10	7
Kuroda3	39.36	6	7.18	4	0.021	2	0.070	4	0.9980	2	115	7	4

*Note:* Harthoorn and van Dalen's method with weights (i.e.,  $g_{ij}$  in (29))  $a_{ij}^2$ ,  $|a_{ij}|$ , and unity are, respectively, denoted by HvD1, HvD/INSD and HvD3. Kuroda method with its corresponding weights in (38), (39), and unity are, respectively, labeled Kuroda1, Kuroda2, and Kuroda3. R. denotes the ranking of the methods, while CmR. is the combined rank of the averages of all the six rankings. 1995  $\Rightarrow$  2000, for example, means that 1995 table is used to project 2000 table. (1995  $\Rightarrow$  2000)  $\Rightarrow$  2005 means that the 2000 updated matrix (on the base of 1995 table) is used as a benchmark for estimating 2005 table. The penalty value is  $M = 10^{100}$ .

the entries that changed their signs (become positive) over time. This is because the EUKLEMS and (G)RAS approaches will always keep original zeros in the updated matrix, but do not generate zero entries themselves. So, 28 zero entries of the 1995 Supply table of the Netherlands become positive in 2000, which effectively represents

the effect of a ‘structural change’ and increased interdependency over this five year interval. The  $N0$  is largest for the ISD method and is equal to 200. This means that the ISD nullifies many elements in its updating process, whose exact number is  $200-28=172$ . These elements would have been negative without introducing the penalty function for (20). Certainly, one would give priority to the method that does *not* produce additional zeros, which are indeed positive in the actual tables.

Besides estimating 2000 table on the basis of the 1995 table (i.e.,  $1995 \Rightarrow 2000$ ), we performed also the following updates: (2) interpolation of the 2000 table on the base of the arithmetic average of the 1995 and 2005 tables,  $0.5 \times (1995 + 2005) \Rightarrow 2000$ , (3) extrapolation of the 2005 table on the basis of the 1995 table,  $1995 \Rightarrow 2005$ , and (4) extrapolation of the 2005 table on the basis of the 2000 estimate that is obtained from the 1995 table,  $(1995 \Rightarrow 2000) \Rightarrow 2005$ . For these four exercises, the count of cells for which  $x_{ij}^{true} = 0$  and  $x_{ij} \neq 0$  were, respectively, equal to 42, 97, 69 and 69 for Supply table projections. The corresponding figures for the Use table projections, which are given in Table 3, were 74, 123, 129 and 129, respectively. Recall that  $MAPE$ ,  $WAPE$ ,  $SWAD$  and  $N0$  do not consider such cases, while they are taken into account by the psi statistic,  $\hat{\psi}$ . From Tables 2 and 3, we observe that the rankings of the updating methods according to  $WAPE$  and  $\hat{\psi}$  measures can be quite similar, especially, for the Supply table projections. This happens for the following reason. If  $x_{ij}^{true} \neq 0$  for all  $i$  and all  $j$ ,  $WAPE/100$  is equal to the *standardized absolute error* (?)<sup>14</sup>

$$SAE = \frac{\sum_{i=1}^m \sum_{j=1}^n |x_{ij} - x_{ij}^{true}|}{\sum_k \sum_l x_{kl}^{true}}.$$

? conclude that “... the versions of information-theoretic statistics favoured in the geographical literature tend to produce essentially the same results as the much simpler  $SAE$ ” (p. 197) by showing that  $\hat{\psi} \approx SAE$  (see their Appendix, pp. 198-200). This also explains the closeness of the values of  $WAPE/100$  and  $\hat{\psi}$  in Tables 2 and 3. Since the number of cases when  $x_{ij}^{true} = 0$  and  $x_{ij} \neq 0$  are less in all four exercises with Supply rather than Use tables projections, the results of  $WAPE$  and  $SAE$  are also closer in the Supply table projections. Thus,  $WAPE$  and  $\hat{\psi}$  provide mainly exactly the same rankings in the Supply estimation (except the last case of  $(1995 \Rightarrow 2000) \Rightarrow 2005$ ), while the rankings are exact in two out of the four projections of the Use tables. Thus we may conclude that, in general, if the number of cells when  $x_{ij}^{true} = 0$  and  $x_{ij} \neq 0$  are relatively small (relative to, say, the

---

<sup>14</sup>In fact, ? use the overall sum of the elements of the *estimated* matrix, but in our setting the last is equal to the overall sum of the true matrix entries, since the row and column sums of the predicted and actual tables are the same.

**Table 3: Results of updating Use tables at purchase prices for the Netherlands**

	MAPE	R.	WAPE	R.	SWAD	R.	$\hat{\psi}$	R.	RSQ	R.	N0	R.	CmR.
1995 $\Rightarrow$ 2000													
EUKLEMS	23.55	5	11.94	6	0.048	6	0.118	6	0.9766	8	72	1	6
RAS	22.58	3	8.97	2	0.031	2	0.089	2	0.9971	1	72	1	2
HvD1	21.93	2	15.08	7	0.068	7	0.149	7	0.9883	6	72	1	5
HvD/INSD	22.82	4	9.07	3	0.031	3	0.090	3	0.997	2	72	1	3
HvD3	266.74	8	30.65	8	0.185	9	0.291	8	0.9778	7	632	9	8
ISD	270.59	9	33.41	9	0.171	8	0.312	9	0.9746	9	950	10	9
IWSD	1086.10	10	37.26	10	0.216	10	0.341	10	0.9684	10	233	7	10
Kuroda1	21.44	1	8.65	1	0.026	1	0.086	1	0.997	2	72	1	1
Kuroda2	62.59	7	11.55	5	0.047	5	0.113	5	0.9949	5	494	8	7
Kuroda3	35.07	6	10.21	4	0.034	4	0.101	4	0.9962	4	163	6	4
0.5 $\times$ (1995 + 2005) $\Rightarrow$ 2000													
EUKLEMS	36.36	5	9.91	6	0.034	6	0.097	7	0.9959	4	18	1	5
RAS	32.34	2	8.14	2	0.023	2	0.080	2	0.9978	9	18	1	2
HvD1	36.02	4	9.85	5	0.035	7	0.097	6	0.9961	5	20	5	6
HvD/INSD	32.11	1	8.10	1	0.024	3	0.080	1	0.9979	10	18	1	1
HvD3	121.86	8	15.44	9	0.071	9	0.146	8	0.9905	2	611	9	8
ISD	133.68	9	16.20	10	0.063	8	0.151	10	0.9922	3	693	10	10
IWSD	222.41	10	15.43	8	0.080	10	0.147	9	0.9864	1	590	8	9
Kuroda1	32.86	3	9.15	3	0.032	5	0.090	4	0.9969	6	18	1	3
Kuroda2	68.72	7	9.97	7	0.030	4	0.096	5	0.9969	6	344	7	7
Kuroda3	43.46	6	9.15	4	0.023	1	0.089	3	0.9972	8	74	6	4
1995 $\Rightarrow$ 2005													
EUKLEMS	53.14	4	22.72	6	0.099	6	0.218	6	0.9686	6	103	1	5
RAS	50.52	3	17.97	1	0.055	1	0.173	2	0.9838	1	103	1	1
HvD1	41.30	1	28.46	7	0.135	7	0.272	7	0.9468	7	103	1	5
HvD/INSD	52.10	5	18.80	3	0.056	3	0.180	3	0.9816	2	103	1	3
HvD3	657.72	9	53.66	8	0.337	9	0.495	8	0.9282	8	873	9	8
ISD	489.10	8	56.82	9	0.310	8	0.516	9	0.9161	9	1244	10	9
IWSD	1988.30	10	62.88	10	0.390	10	0.561	10	0.8915	10	286	7	10
Kuroda1	47.52	2	18.07	2	0.056	2	0.172	1	0.9789	4	103	1	2
Kuroda2	127.56	7	22.62	5	0.092	5	0.214	5	0.9779	5	778	8	7
Kuroda3	74.87	6	20.19	4	0.060	4	0.192	4	0.9796	3	234	6	4
(1995 $\Rightarrow$ 2000) $\Rightarrow$ 2005													
EUKLEMS	53.14	5	22.72	5	0.099	6	0.218	5	0.9686	6	103	1	5
RAS	50.52	3	17.97	1	0.055	1	0.173	1	0.9838	1	103	1	1
HvD1	41.09	1	30.83	7	0.149	7	0.295	7	0.9313	7	103	1	6
HvD/INSD	51.26	4	18.60	2	0.056	2	0.178	2	0.982	2	103	1	2
HvD3	887.95	9	56.97	8	0.334	9	0.577	8	0.9254	8	398	8	8
ISD	795.49	8	60.43	9	0.315	8	0.622	9	0.9147	9	521	10	9
IWSD	3281.60	10	72.50	10	0.390	10	1.139	10	0.5765	10	186	6	10
Kuroda1	47.87	2	18.87	3	0.061	3	0.180	3	0.9727	5	103	1	3
Kuroda2	187.77	7	23.10	6	0.091	5	0.233	6	0.9781	4	476	9	7
Kuroda3	80.82	6	20.27	4	0.061	3	0.195	4	0.9798	3	200	7	4

*Note:* See the endnotes to Table 2.

size of the projected matrix), then  $WAPE$ ,  $SAE$  and  $\hat{\psi}$  are essentially equivalent goodness-of-fit statistics for matrix comparative purposes.

In all parts of Table 2 and Table 3 we systematically observe that three methods HvD3, ISD and IWSD are overall performing worse among all ten alternatives, and produce deviations of the values of the statistics that are far away from those of the other methods. They share one common feature: the errors (differences) of the new over old elements ratios,  $x_{ij}/a_{ij}$ , from unity in the objective functions are weighted

directly by the *enlarged* size of the original elements (see (20), (25), (29) with  $g_{ij} = 1$  (HvD3)). This suggests that a weighting scheme that weights deviations of the ratios by the expanded size of the original elements (i.e., taking the element to the power of more than one) may not be a good starting point. Given the poor performance of these three approaches, our first conclusion is that these methods should *not* be used for updating SUTs (and SIOTs).

All parts of Table 2 and Table 3 (except the case  $0.5 \times (1995 + 2005) \Rightarrow 2000$  in the Supply projection) consistently show that the three ‘best’ performing updating methods are RAS, HvD/INSD,<sup>15</sup> and Kuroda1. The overall ranking of the four exercises in Table 2 and also in Table 3 show the following ordering: RAS, Kuroda1 and HvD/INSD. The fact that (G)RAS was outperforming other updating methods was extensively reported in the literature (see e.g., Evans and Lindley 1973, Davis et al. 1977, Jalili 2000, Jackson and Murray 2004). Huang et al. (2008) recommended to use GRAS and INSD for updating IO tables, and our results confirm that indeed these two methods are not inferior. However, there is, at least, one other method which is performing quite well either. This is Kuroda’s method that minimizes the changes of the updated row and column *coefficients* from those of the base year (see (40)). Note that HvD/INSD has the objective that weights the corresponding updating errors of the transactions’ *ratios* directly by the size of the original transactions  $|a_{ij}|$ . The same is true for (G)RAS (see (11)). This would suggest that the hypothesis that transactions’ *ratios* (and not transactions themselves) are more permanent and stable for larger original deliveries than for the smaller ones is indirectly confirmed by the success of RAS and HvD/INSD, at least, for the case of the Netherlands. Our second conclusion is that, for our data, the RAS, Kuroda1, and HvD/INSD updating methods are performing quite well, and may be used for estimating SUTs. The relative performance of these methods should be further checked on other datasets as well.

Comparing the results in the first and second blocks of Tables 2 and 3 we can conclude that the interpolation of the 2000 Supply table does *not* provide better estimate than its extrapolation on the basis of the 1995 Supply table according to all methods except the three worse performing updating approaches (i.e., HvD3, ISD and IWSD). These methods turn out to give better results when both 1995 and 2005 tables are used, but still the obtained errors are quite large compared to those of the other approaches. However, in the estimation of the Use table, the interpolation may give better estimates than pure extrapolation on the earlier data. Since, the Use table is much denser than the Supply table, one may conclude that the interpolation may give better projections for the tables that are not sparse.

---

<sup>15</sup>Recall that Harthoorn and van Dalen’s method with the elements’ relative confidences equal to the absolute values of the original entries ( $g_{ij} = |a_{ij}|$  in (29)) is nothing else as the INSD approach.

The bottom parts of Table 2 and Table 3 give the result of projecting 2005 SUTs using the updated 2000 tables, which were based on the 1995 SUTs. The question that we are interested in is whether using the new 2000 estimate gives better outcome than the earlier 1995 actual data for updating 2005 tables. Comparing the leading four updating methods for all indicators we get no definitive answer. Kuroda1 method mainly performs worse with the recent estimate rather than the actual earlier data. HvD/INSD's performance can go either directions depending on the choice of the goodness-of-fit statistic. Note that the EUKLEMS and RAS give exactly the same values of *all* the indicators in the current exercise. The reason is simple: for these methods the updating case (1995  $\Rightarrow$  2000)  $\Rightarrow$  2005 means nothing else as using the structure of the 1995 SUTs to predict the 2005 matrices. Thus the results of their performance in this updating exercise will be exactly the same as the estimation of the 2005 tables directly from the 1995 data.

In our final empirical application, we use Spanish SUTs for two benchmark years of 2000 and 2005 available from the National Statistics Institute of Spain. The distinguishing feature of this data is that the Use tables are given in basic prices, and further distinguished between the domestic and imported use matrices. This is exactly what we need for assessing the Euro method that has not been analyzed in the previous evaluations namely because of unavailability of the Netherlands' domestic and imported Use tables in basic prices. The Spanish Supply and intermediate Use tables originally consisted of 118 commodities and 75 industries. After some 'cleaning process' and the fact that the Euro method works only for symmetric SUTs, we ended up with 73 products and 73 industries.<sup>16</sup> The domestic and imported final demand consist of three categories: total consumption, gross capital formation (with both positive and negative entries), and exports.

The results of the evaluation of seven methods, six of which performed relatively well with the Dutch SUTs, are given in Table 4. We do not compute the coefficient of determination this time, since in the previous exercises it showed to be a weak statistic for matrix comparison purposes. In fact, such an outcome confirmed the same conclusion of Knudsen and Fotheringham (1986) on the use of *RSQ*. To jump ahead, the analysis of Table 4 reveals that the same three methods always outperform other updating techniques, and, moreover, their *overall* ordering is also similar: GRAS, Kuroda1, and HvD/INSD. Note that in the prediction the total intermediate use Kuroda1 outperforms RAS. These additional findings show that indeed the three mentioned methods can be used for updating SUTs and SIOTs.

---

<sup>16</sup>This number decreased from 75 to 73 because we were not able to distinguish between Market and Non-market health services, and Market and Non-market recreational, cultural and sporting activities. This is because CNPA 96 products classification does not distinguish between non-profit institution serving households (NPISHs) and governmental services, which is the case for CNAE 93 industry classification.

**Table 4: Results of updating Spanish SUTs at basic prices**

	MAPE	R.	WAPE	R.	SWAD	R.	$\hat{\psi}$	R.	N0	R.	CmR.
Supply Table (2000 $\Rightarrow$ 2005)											
EURO	21.04	6	8.44	7	0.109	7	0.083	6	108	1	7
EUKLEMS	18.91	3	2.51	5	0.004	3	0.024	4	108	1	4
RAS	19.61	4	1.96	1	0.004	2	0.018	1	108	1	1
HvD1	16.43	1	2.65	6	0.010	6	0.025	5	109	2	5
HvD/INSD	19.88	5	2.05	3	0.004	1	0.019	2	115	4	3
Kuroda1	18.40	2	1.98	2	0.005	4	0.018	1	110	3	2
Kuroda3	27.49	7	2.30	4	0.005	5	0.022	3	183	5	6
Total Intermediate Use Table (2000 $\Rightarrow$ 2005)											
EURO	38.02	4	20.93	7	0.348	8	0.206	7	120	2	7
EUKLEMS	93.59	7	22.03	8	0.195	7	0.214	8	112	1	8
EUKLEMS*	34.68	1	16.58	5	0.150	6	0.163	5	120	2	3
RAS	38.07	5	13.81	2	0.077	2	0.136	2	120	2	2
HvD1	35.43	2	20.66	6	0.083	4	0.203	6	120	2	5
HvD/INSD	39.15	6	15.02	3	0.099	5	0.148	3	120	2	3
Kuroda1	37.69	3	13.17	1	0.023	1	0.129	1	120	2	1
Kuroda3	96.12	8	15.75	4	0.083	3	0.154	4	364	3	6
Total Final Demand Table (2000 $\Rightarrow$ 2005)											
EURO	249.42	7	8.38	7	0.051	7	0.085	7	3	1	7
EUKLEMS	235.18	5	14.78	8	0.100	8	0.148	8	3	1	8
EUKLEMS*	239.91	6	7.90	6	0.048	6	0.080	6	3	1	6
GRAS	114.35	2	4.28	2	0.012	2	0.044	1	3	1	1
HvD1	143.88	3	5.53	5	0.019	5	0.056	5	4	2	4
HvD/INSD	90.04	1	4.30	3	0.012	3	0.044	3	5	3	3
Kuroda1	227.74	4	4.24	1	0.010	1	0.044	1	4	2	2
Kuroda3	269.61	8	4.56	4	0.017	4	0.047	4	6	4	5

*Note:* *EUKLEMS\** technique differs from the *EUKLEMS* described in Section 2.1 in only projecting the Use matrix on the base of the Use table at basic prices (and not purchase prices), where the growth rates of commodity outputs at basic prices are applied row-wise to the original Use table and normalized such that the actual totals are consistent with the true use totals (similar to Step 4 in Section 2.1).

We consider one more indicator in case of the Supply estimation, which has to do with the change in the size of the main product/industry compared to its non-principal activities and by-products.<sup>17</sup> Let  $s_{ij}$  be the amount of product  $i$  that is supplied by industry  $j$ . Then the main product supplying ratio of commodity  $i$  is

$$MPSR_i = \frac{s_{ii}}{\sum_j s_{ij}},$$

while the main industry supplying ratio of industry  $j$  is

$$MISR_j = \frac{s_{jj}}{\sum_i s_{ij}}.$$

<sup>17</sup>We are grateful to Jiemin Guo for this suggestion.

To see whether the estimated Supply retains the same main product and industry ratios as in the true Supply matrix, we compute the product and industry difference indices for commodity  $i$  and industry  $j$ , respectively, as

$$PD_i = MPSR_i / \widehat{MPSR}_i \quad \text{and} \quad ID_j = MISR_j / \widehat{MISR}_j, \quad (41)$$

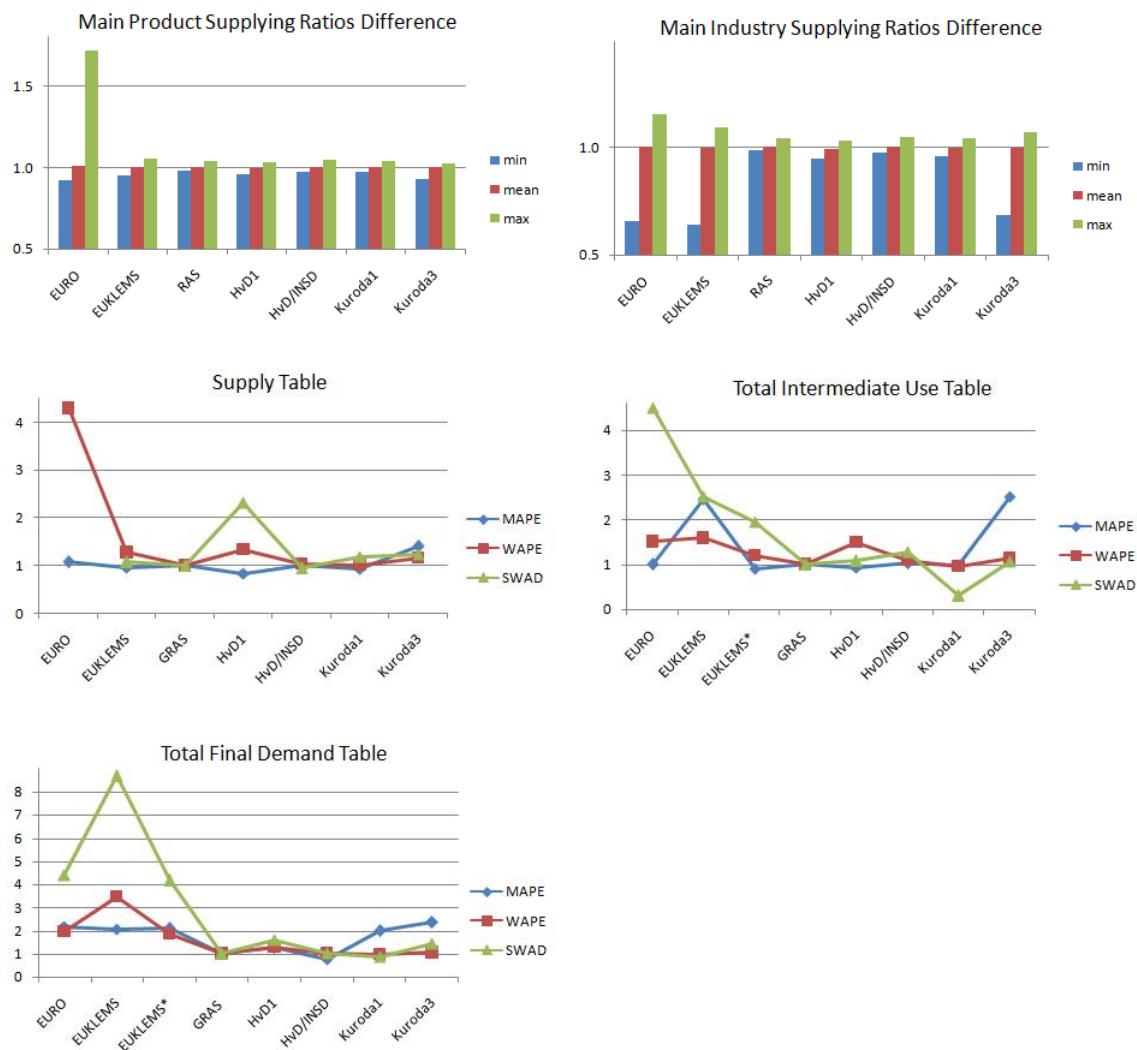
where, for example,  $\widehat{MPSR}_i$  is the MPSR of product  $i$  in the *estimated* Supply matrix. The upper part of Figure 1 illustrates the descriptive statistics (i.e., minimum, mean and maximum) of these indicators. It shows that although the average of the product and industry difference indices (41) of all methods is pretty close to one, the spread of these indices across products and industries largely differs from one method to the other. In particular, we clearly see that the Euro and EUKLEMS methods are performing worse in this regard (although EUKLEMS approach shows better performance with MPSRs differences), while RAS, HvD1, HvD/INSD, and Kuroda1 show quite small deviation of both difference indicators. This clearly suggests that the main product and industry supplying ratio are predicted quite inaccurately across commodities and sectors by the Euro, EUKLEMS, and Kuroda3 methods.

The bottom part of Figure 1 illustrates *relative MAPE*, *WAPE* and *SWAD* indicators derived from Table 4 for Supply, total intermediate and final use projections. That is, these indicators are normalized by corresponding values of the GRAS method, hence for GRAS these relative measures are equal to one. We do not show  $\hat{\psi}$  statistic, because its trend is extremely close to *WAPE*. We do not show the relative SWAD measure for the Supply projection for the Euro method, since this number is quite large (i.e., equal to 26.63) that makes the visual comparison of other methods impossible. These three figures clearly demonstrate our main finding: the relative measures for HvD/INSD and Kuroda1 are very close to one (i.e., to the GRAS outcome), and those of all other methods are mainly larger than unity, thus performing worse than GRAS. The fourth overall ranking takes the EUKLEMS method in predicting Supply tables, but it estimates the Use tables with large errors (together with the Euro method).

EUKLEMS\* differs from the EUKLEMS described in Section 2.1 in only one respect: it projects the Use matrix on the base of the Use table at basic prices (and not purchasers' prices), where the growth rates of commodity outputs at basic prices are applied row-wise to the original intermediate and final use table and then normalized such that the actual totals are consistent with the true use totals of the 2005 Use tables (similar to Step 4 in Section 2.1). This procedure show better result than the standard EUKLEMS, where the stating point is the Use table at purchasers' prices. This result is entirely expectable, since given that the Use table



Figure 1: Results of Spanish SUTs projections



at basic prices is available, one does not have to transform Use tables from purchase into basic prices. Hence, additional errors due to the estimation of matrices of trade margins, transportation margins, and net taxes are avoided when the initial Use table is available at basic prices.

## 4 Concluding remarks

Given the importance of availability of Supply and Use tables (SUTs) and symmetric input-output tables (SIOTs) for policy-relevant research and the problem of timeliness of these data, many non-survey (semi-survey) methods have been proposed and extensively used in the literature. The purpose of this paper was to present and assess the relative performance of eight existing methods for updating SUTs

(and SIOTs), some of which are less (or even not) known in the literature. Three methods were slightly revised in this paper with regard to updating negative entries as well as preserving the signs of the elements of the original matrix in the projected one.

Our empirical applications of the methods to the Dutch and Spanish SUTs projections showed that (G)RAS, and two other methods proposed, respectively, by Harthoorn and van Dalen (1987) and Kuroda (1988) are providing best estimates of the SUTs of the Netherlands and Spain. Hence, we conclude that these three methods may be used for updating purposes. We should note that the improved normalized squared differences (INSD) updating approach proposed by Huang et al. (2008) is nothing else as a particular case of Harthoorn and van Dalen's method. On the other hand, among other techniques, the Euro method that is used by the Eurostat does not show satisfying performance on these data.

Our study also suggests that the hypothesis that transactions' *ratios* (and not transactions themselves) are more permanent and stable for larger original deliveries than for the smaller ones is indirectly confirmed by the relative success of GRAS and HvD/INSD methods. Finally, when there is a choice between extrapolation on the base of earlier tables and interpolation that also uses more recent information, we could conclude from the results that interpolation, on average, works better for denser matrices (such as Use tables), while it is outperformed by extrapolation results for sparse matrices (such as Supply tables). We think that the reason for this outcome is as follows. Over time, the degree of interconnectedness becomes larger, and this is especially true for Use tables where more zero elements become non-zero (mainly positive). Thus, extrapolation on earlier data will miss many such changes, while if one uses an earlier and more recent tables, it is expected that such "structural changes" are taken into account more accurately.

## References

- Allen, R. and W. Gossling (eds.): 1975, *Estimating and Projecting Input-Output Coefficients*. London: Input-Output Publishing Company.
- Almon, C.: 1968, 'Recent methodological advances in input-output in the United States and Canada'. Paper presented at the Forth International Conference on Input-Output Techniques in Geneva.
- Bacharach, M.: 1970, *Biproportional Matrices and Input-Output Change*. Cambridge: Cambridge University Press.
- Beutel, J.: 2002, 'The economic impact of objective 1 interventions for the period 2000-2006'. Report to the Directorate-General for Regional Policies, Konstanz.
- Butterfield, M. and T. Mules: 1980, 'A testing routine fo evaluating cell by cell accuracy in short-cut regional input-output tables'. *Journal of Regional Science* **20**, 293–310.
- Carey, M., C. Hendrickson, and K. Siddiharthan: 1981, 'A method for direct estimation of origin/destination trip matrices'. *Transportation Science* **15**, 32–49.
- Davis, H. C., E. M. Lofting, and J. A. Sathaye: 1977, 'A comparison of alternative methods of updating input-output coefficients'. *Technological Forecasting and Social Change* **10**, 79–87.
- de Mesnard, L. and R. E. Miller: 2006, 'A note on added information in the RAS procedure: reexamination of some evidence'. *Journal of Regional Science* **46**, 517–528.
- Eurostat: 2008, *European Manual of Supply, Use and Input-Output Tables. Methodologies and Working Papers*. Luxembourg: Office for Official Publications of the European Communities.
- Evans, G. J. and R. M. Lindley: 1973, 'The use of RAS and related models in manpower forecasting'. *Economics of Planning* **13**, 53–73.
- Friedlander, D.: 1961, 'A technique for estimating a contingency table, given marginal totals and some supplemental data'. *Journal of the Royal Statitistical Society, Series A* **124**, 412–420.
- Gilchrist, D. A. and L. V. St. Louis: 1999, 'Completing input-output tables using partial information, with an application to Canadian data'. *Economic Systems Research* **11**, 185–194.
- Golan, A., G. Judge, and S. Robinson: 1994, 'Information from incomplete or partial multisectoral economic data'. *Review of Economics and Statistics* **76**, 541–549.

- Golan, A. and S. J. Vogel: 2000, 'Estimation of non-stationary social accounting matrix coefficients with supply-side information'. *Economic Systems Research* **12**, 447–471.
- Günlük-Şenesen, G. and J. M. Bates: 1988, 'Some experiments with methods of adjusting unbalanced data matrices'. *Journal of the Royal Statistical Society, Series A* **151**, 473–490.
- Harrigan, F. and I. Buchanan: 1984, 'A quadratic programming approach to input-output estimation and simulation'. *Journal of Regional Science* **24**, 339–358.
- Harthoorn, R. and J. van Dalen: 1987, 'On the adjustment of tables with Lagrange multipliers'. **NA-024**. Central Bureau of Statistics, The Netherlands, National Accounts Research Division.
- Hinojosa, R. C.: 1978, 'A performance test of the biproportional adjustment of input-output coefficients'. *Environment and Planning A* **10**, 1047–1052.
- Huang, W., S. Kobayashi, and H. Tanji: 2008, 'Updating an input-output matrix with sign-preservation: some improved objective functions and their solutions'. *Economic Systems Research* **20**, 111–123.
- Jackson, R. W. and A. T. Murray: 2004, 'Alternative input-output matrix updating formulations'. *Economic Systems Research* **16**, 135–148.
- Jalili, A. R.: 2000, 'Evaluating relative performances of four non-survey techniques of updating input-output coefficients'. *Economics and Planning* **33**, 221–237.
- Johnston, R. and C. Pattie: 1993, 'Entropy-maximizing and the iterative proportional fitting procedure'. *Professional Geographer* **45**, 317–322.
- Junius, T. and J. Oosterhaven: 2003, 'The solution of updating or regionalizing a matrix with both positive and negative entries'. *Economic Systems Research* **15**, 87–96.
- Knudsen, D. C. and A. S. Fotheringham: 1986, 'Matrix comparison, goodness-of-fit, and spatial interaction modeling'. *International Regional Science Review* **10**, 127–147.
- Kullback, S.: 1959, *Information Theory and Statistics*. New York: John Wiley and Sons.
- Kuroda, M.: 1988, 'A method of estimation for the updating transaction matrix in the input-output relationships'. In: K. Uno and S. Shishido (eds.): *Statistical Data Bank Systems. Socio-economic database and model building in Japan*. Amsterdam: North Holland, pp. 43–56.
- Lahr, M. L.: 2001, 'A strategy for producing hybrid regional input-output tables'. In: M. L. Lahr and E. Dietzenbacher (eds.): *Input-Output Analysis: Frontiers and Extensions*. New York: Palgrave, pp. 211–242.

- Lahr, M. L. and L. de Mesnard: 2004, 'Biproportional techniques in input-output analysis: table updating and structural analysis'. *Economic Systems Research* **16**, 115–134.
- Lecomber, J.: 1975, 'A critique of methods of adjusting, updating and projecting matrices'. In: R. Allen and W. Gossling (eds.): *Estimating and Projecting Input-Output Coefficients*. London: Input-Output Publishing Company, pp. 43–56.
- Lenzen, M., R. Wood, and B. Gallego: 2007, 'Some comments on the GRAS method'. *Economic Systems Research* **19**, 461–465.
- Leontief, W. W.: 1941, *The Structure of American Economy, 1919-1929: An Empirical Application of Equilibrium Analysis*. Cambridge: Cambridge University Press.
- Mari Bhat, P.: 1990, 'Estimating transition probabilities of age misstatement'. *Demography* **27**, 149–163.
- McDougall, R. A.: 1999, 'Entropy theory and RAS are friends'. GTAP Working Paper no. 06.
- Miller, R. E. and P. D. Blair: 2009, *Input-Output Analysis: Foundations and Extensions*. Cambridge: Cambridge University Press, 2nd edition.
- Mínguez, R., J. Oosterhaven, and F. Escobedo: 2009, 'Cell-corrected RAS method (CRAS) for updating or regionalizing an input-output matrix'. *Journal of Regional Science* **49**, 329–348.
- Parikh, A.: 1979, 'Forecasts of input-output matrices using the RAS method'. *Review of Economics and Statistics* **61**, 477–481.
- Pavia, J. M., B. Cabrer, and R. Sala: 2009, 'Updating input-output matrices: assessing alternatives through simulation'. *Journal of Statistical Computation and Simulation* **79**, 1467–1482.
- Rees, P.: 1979, 'Regional population projection models and accounting methods'. *Journal of the Royal Statistical Society, Series A* **142**, 223–255.
- Schoen, R. and S. Jonsson: 2003, 'Estimating multistage transition rates from population distributions'. *Demographic Research* **9**, 1–24.
- Snower, D. J.: 1990, 'New methods of updating input-output matrices'. *Economic Systems Research* **2**, 27–38.
- St Louis, L. V.: 1989, 'Empirical tests of some semi-survey update procedures applied to rectangular input-output tables'. *Journal of Regional Science* **29**, 373–385.
- Stone, R.: 1961, *Input-Output and National Accounts*. Paris: Organization for European Economic Cooperation.

- Strømman, A. H.: 2009, 'A multi-objective assessment of input-output matrix updating methods'. *Economic Systems Research* **21**, 81–88.
- Tilanus, C. B.: 1968, *Input-Output Experiments: the Netherlands, 1948-1961*. Rotterdam: Rotterdam University Press.
- Timmer, M. P., P. Aulin-Ahmavaara, and M. Ho: 2005, *EUKLEMS Road map WP1*. available at: [www.euklems.net/workpackages/roadmap-wp1\\_12-10-2005.pdf](http://www.euklems.net/workpackages/roadmap-wp1_12-10-2005.pdf).
- Wilcoxon, P. J.: 1989, 'Kuroda's method for constructing consistent input-output data sets'. Impact Research Centre, University of Melbourne.