UB economics WORKING PAPERS
Col.lecció d'Economia E17/365

Technology of the preferences: linking consumption expenditures to value added with minimal information

Esteban Fernández-Vázquez
Mònica Serrano

UNIVERSITAT DE BARCELONA

Technology of the preferences: linking consumption expenditures to value added with minimal information

Abstract: The combination of consumption data from expenditure surveys with information contained in IO tables is a crucial step to structural change analysis. Herrendorf et al. (2013) focus the attention on applying a consistent definition of commodities on both the household and production sides —i.e. estimation of utility and production functions— when connecting models with data in any multisector general equilibrium model. The point of departure of these analyses consist, basically, on connecting the information on consumption made by households with the final demand vector (or matrix) present in the IO tables, which is then conveniently modified to produce the multipliers of interest. This process requires the construction of a concordance or bridge matrix to make this connection possible, since several issues affect the combination of these two data sources: differences in price valuation between consumption surveys and IO tables, the influence of taxes and margins or the different product classifications between these two frameworks make this combination a challenge for the researcher.

In this paper we explore this challenge with a twofold purpose: (i) to investigate how important a “good” or “bad” conciliation of our consumption data between household surveys and IO tables affect our results in terms of the so-called total requirement matrix or impact analysis; and (ii) to propose a conciliation technique between both data structure, which using only minimal information provides a systematic way or reconciling them if detailed data are not at hand. This technique is based on entropy econometrics and it allows making statistical inference on the bridge matrix estimated. Both research objectives are illustrated by means of numerical simulation and by its application to a real-world case.

JEL Codes: C1, O11,O41.

Keywords: Expenditure surveys, Input-output tables, Structural change, Multisector model, Econometrics.

Esteban Fernández-Vázquez
Universidad de Oviedo

Mònica Serrano
Universitat de Barcelona

ISSN 1136-8365

1. Introduction

Kuznets (1966, 1973) determined structural change, defined as the reallocation of economic activity across the three broad sectors —i.e. agriculture, manufacturing and services—, as one of the six stylized facts of economic growth and development. As economies develop behaviour of sectoral aggregate variables, such value added or growth rate, changes: agriculture diminishes, manufacturing increases and then diminishes, and services increases.

Recent literature has analysed the driving forces that determines structural change. There are two theoretical mechanisms that links the sector specific household expenditure structure across sectors to structural change. On the one hand, structural change is driven by income effects that are generated by non-homothetic preferences when income changes but relative prices remains the same; in this case, technological progress is assumed to be uniform across all sectors (Kongsamut et al., 2001). On the other hand, structural change is driven only by changes in relative prices and necessarily technological progress among sectors is different (Baumol, 1967; Nagai and Pissarides, 2007). Although empirical evidence shows that both income and relative prices changed over the past, there remains no consensus about the relative importance of the two main determinants as drivers of structural change.

Whereas Herrendorf et al. (2015) analysed the importance of differential rates of technological progress among the three consumption sectors, Herrendorf et al. (2013) focused on the preference side of structural change —i.e. they analysed the relative importance of both changes in income and changes in relative prices in the households' consumption bundles as a source of structural change— showing which preference specification should be used in applied work on structural change. Under the so-called consumption value-added approach changes in income are less important than changes in relative prices being a homothetic Leontief utility function the specification that provides a good fit to the US time series data (1947-2010). Alternatively, a non-homothetic Stone-Gary utility function provides a better result under the so-called final consumption expenditure approach since changes in income rather than changes in relative prices are the dominant force behind changes in expenditures shares.

This result not only provides an estimation of the utility functions and asses the implications for the determinants of the structural change, but also contributes to clarify an essential ambiguity about how to define commodities when the research needs to link a multisector general equilibrium model to the empirical data.[i] At this point the attention focuses on applying a consistent definition of commodities on both the household side when we estimate utility functions —i.e. preferences specification— and the production side when we estimate production functions —i.e. technology specification—. Whereas in multisector models the sector classification is the same no matter which agent is using or producing them —i.e. the three broad sectors agriculture, manufacturing and services are the same for the consumption and production side—, the data show a different reality.

In national accounts both activities —consumption and production— are classified following different criteria and generally recorded in different data sources. On the one hand, following the System of National Accounts (SNA) commodities purchased and consumed by households are classified attending at the purpose or function of such commodity. This is known as the *Classification of individual consumption by purpose*

(COICOP). On the other hand, the classification of the production of such commodities follows the so-called International standard industrial classification of all economic activities (ISIC).[ii] That is household buy a medicine, but do not directly buy a chemical product produced by the chemical industry. In the particular case of USA both classification are the *New personal consumption expenditures* (NPCE)[iii] by function and the *North American industry classification system* (NAIC).[iv] Regarding data sources, consumption data commonly come from national household income and expenditure surveys, whereas production data come from the input-output (IO) benchmark of the national accounts.

In addition to this different classification generally there are other differences that worth to be mentioned. First, in sector models each sector produces only one product, whereas in data sectors produce more than one product. Second, production data is in basic prices —or producer's prices in USA—, whereas consumption data is in purchaser's prices. Third, production data provides information about the total supply of commodities in the economy distinguishing domestic production and imports; however, consumption data only informs about total expenditure without distinguishing which share of imports is consumed. Finally, when working with final consumption expenditure time series data, final consumption quantities should be measured using chain-weighted indices. All together makes the construction of multisector models from data far to be an obvious task. The proper connection may have significant effects not only on the analysis of structural change but also in the comparison across studies.

The process to solve this challenge requires the construction of a concordance or bridge matrix to make this connection possible. However, national bridge matrices are not available for all countries or only partial information is offered it the best cases. Therefore, the bridge matrix should be estimated, which it is a challenge for the researcher taking into account all the issues affect the combination of these two data sources mentioned above.

Herrendorf et al. (2013) is one of the first studies that become aware of the importance of the consistency between sector models and data within modern economic literature focused on structural change literature. They constructed final expenditure in producer's prices —removing distribution cost—, linked consumption expenditures to value added —using the total requirement matrix and the industry-by-commodity total requirement matrix—, and obtained final consumption quantities —applying chain-weighted indices—. In addition, they approximated the connection between final consumption expenditure data and consumption value added data. However, the specific intricate input-output relationships between both data structures remains an unsolved challenge.

The proper specification of these relationships —the so-called bridge matrix— is an important issue since they implicitly translate part of the income effects that dominate with final consumption expenditure into relative prices effects that are much more important with consumption value-added, and vice versa. In other words, the bridge matrix represents a technology that combines intermediate goods (produced by industrial sectors) to final goods that are consumed by households. The intrinsic characteristics of this technology differ from the traditional technology matrix that represent the combination of intermediate input, labour and capital to produce intermediate inputs. In that sense we called the technology represented by the bridge matrix "technology of the preferences".

In this paper we explore this challenge with a twofold purpose: (i) to investigate how important a “good” or “bad” conciliation of our consumption data between consumption data —i.e. household income and expenditure surveys— and production data —i.e. IO tables— affect our results in terms of the so-called total requirement matrix or impact analysis; and (ii) to propose a conciliation technique between both data structure, which using only minimal information provides a systematic way or reconciling them if detailed data are not at hand. This technique is based on entropy econometrics and it allows making statistical inference on the bridge matrix estimated. Both research objectives are illustrated by means of numerical simulation and by its application to real-world cases.

Pursuing these objective, this paper offers two contributions. First, we contribute to the structural change literature by providing an assessment of the importance of the correct specification of the bridge matrix. Second, we provide a technique that allows to obtain such bridge matrix when the necessary detailed data is not available.

An outline of the paper follows. In the next section we explore the role played by the bridge matrix on the estimation of the total requirement matrix or impact analysis. In section 3 we study the importance of the correct specification of the elements of the bridge matrix. Section 4 explores the possibilities of estimating the elements of the bridge matrix when only partial or minimal information is available for the researcher. Section 5 concludes.

2. Linking expenditure and production data: expenditure surveys and IO tables

On this section we explore the role played by the bridge matrix (**B**, hereafter) on the estimation of impacts of final private consumption on a multisector model as the input-output model. Obviously, all the conclusions can be easily applied to the total requirement matrix used by Herrendorf et al. (2013) in their analysis of structural change. Our point of departure is a matrix **B** with dimensions $(n \times p)$, being n the number of industries in the IO table that will be the base of our model and p the number of product categories that can be identified on a household consumption survey. In practical terms, n is usually set by a CPC classification while p is determined following the COICOP coding. A typical cell of matrix **B**, b_{ik}, measures how much of the private consumption in product k should be attributed to industry i. Consequently, **B** is formed by columns that sum up to one (i.e., $\sum_{i=1}^{n} b_{ik} = 1; k = 1,..p$) and it can be interpreted as a matrix describing the technology that relates the consumption patterns of the household with the production of the industries.

Defining **c** as the $(p \times 1)$ vector of final private consumption for the different p products considered on the household surveys, and $\mathbf{y^c}$ the $(n \times 1)$ vector of private consumption included in the IO tables -as part of the final demand vector **y** for a model with n industries considered-, we can write:[1]

[1] Note that $\sum_{i=1}^{n} y_i^c = \sum_{k=1}^{p} c_k$; i.e., the aggregate final consumption in the economy is the same across types of consumption products or across industries.

$$\mathbf{y^c} = \mathbf{Bc} \tag{1}$$

This equation can be easily connected with the standard quantity input-output model:

$$\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{y} = \mathbf{Ly} \tag{2}$$

Being $\mathbf{A}$ the $(n \times n)$ matrix of technical coefficients, $\mathbf{I}$ the identity matrix of a proper dimension, $\mathbf{x}$ the $(n \times 1)$ vector of output per industry, and $\mathbf{L}$ the so-called total requirement matrix (Herrendorf et al., 2013) or inverse Leontief matrix. The output generated by the private final consumption ($\mathbf{x^c}$) can be written as:

$$\mathbf{x^c} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{y^c} = \mathbf{Ly^c} = \mathbf{LBc} \tag{3}$$

Equation (3) explains how changes in the private consumption $\mathbf{c}$, as we have it classified on household surveys, are transformed into changes in the output at the industry level at the industry classification we have in our input-output model. Impact analysis of changes in consumption patterns is usually conducted by applying the following stages:

1. plausible levels of $\mathbf{c}$ under a scenario of interest are assumed or estimated.
2. values on $\mathbf{c}$ are converted into values on $\mathbf{y^c}$. This step requires the specification of a matrix **B** in order to apply equation (1).
3. a vector of industry output derived from the new consumption levels $\mathbf{x^c}$ are calculated by applying equation (3).

When conducting this type of impact analysis, input-output researchers usually pay most of their attention to having accurate estimates of the elements of the inverse Leontief matrix **L**. The availability of IO tables in order to derive the cells on **A** and **L** matrices is often considered as the crucial step in the empirical analysis. Little attention is, on the contrary, paid to the specification of matrix **B**: most of the empirical studies that estimate impacts of private consumption by applying an IO model do not provide details on the explicit or implicit specification of **B**. The concordance between the classification observable on the household consumption surveys and the industry classification is usually an *ad-hoc* process based on the "subjective" similarity that the researcher can detect between both classifications.

This problem is mainly caused by the absence of a regular series of "objective" matrices estimated by the statistical agencies. With some exceptions, the statistical agencies that publish the IO benchmark on a regular basis do not make available the bridge matrices that the researcher would require to perform impact analysis of consumption changes. Researchers could base on some qualitative mapping at a very disaggregated level[2] but, even so, the usual level of aggregation on which consumption data are reported on expenditure surveys does not allow for a perfect identification of the industries to which consumption levels should be assigned. In other words, the researcher could know that the expenditure on product k should be attributed to the production of industries i and j,

[2] See, for example, the concordance table published by Eurostat for linking the CPA 2002 with the COICOP 1999 classifications at: http://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL&StrLanguageCode=ES&IntCurrentPage=4.

but in absence of an official **B** matrix, proportions b_{ik} and b_{jk} cannot be exactly known but somehow arbitrarily set. The following section attempts to measure the consequences of the specification on the elements of **B** when this matrix is not available.

3. How important is B? A numerical experiment

This section explore the role played by the specification of the b_{ik} elements on impact studies of consumption by an input-output model. In order to do that, we depart from an observable and assumed as "true" **B** matrix that links expenditure data on household surveys **c** with the vector of final private consumption on an input-output model $\mathbf{y^c}$ by means of equation (1). In particular, we set our **B** matrix from the data published by Statistics Denmark on IO data for 2015, which releases a matrix that relates consumption data for 41 types of products following the COICOP classification (*p*) with the 68 industries (*n*) contained in the IO benchmark.[3] More specifically, we focus our interest on the data linking household consumption only with Danish production.[4] The figures on this matrix have been conveniently scaled to sum up to one by column. Since **c**, **B** and $\mathbf{y^c}$ are observable, the estimation of the output generated by the final private consumption in the Danish economy is immediate by applying equation (3).

In order to quantify the effect of the elements of matrix **B** on the impacts of consumption, we have simulated alternative bridge matrices $\widetilde{\mathbf{B}}$, which are assumed to be the specification of the true **B** in a situation where it is not directly observable. Once different $\widetilde{\mathbf{B}}$ matrices are simulated, and keeping constant the figures on vector **c** and matrix **L**, we have calculated the vector of final private consumption by industry that would have obtained ($\tilde{\mathbf{y}}^\mathbf{c}$) and the corresponding output by industry ($\tilde{\mathbf{x}}^\mathbf{c}$) by means of the expression:

$$\tilde{\mathbf{x}}^\mathbf{c} = \mathbf{L}\tilde{\mathbf{y}}^\mathbf{c} = \mathbf{L}\,\widetilde{\mathbf{B}}\mathbf{c} \tag{4}$$

The elements $\tilde{b}_{ik}$ are simulated from the true b_{ik} as:

$$\tilde{b}_{ik} = b_{ik} + \varepsilon_{ik} \tag{5}$$

Where $\varepsilon_{ik} \sim N(0, \delta b_{ik})$. This implies that the elements in $\widetilde{\mathbf{B}}$ are assumed to deviate from the true elements with zero mean and a standard deviation that is proportional to the size of the original element b_{ik}. The scalar δ sets the scale of the standard deviation and is an indicator of the size of the deviations between the true cells b_{ik} and the $\tilde{b}_{ik}$ actually applied in equation (4) when calculating the consumption impacts. Note that generating the $\tilde{b}_{ik}$ elements as in (5) keeps unaltered the qualitative mapping between the n industries and the p products as it is on **B**; i.e., null elements b_{ik} generate null $\tilde{b}_{ik}$ cells for matrix $\widetilde{\mathbf{B}}$ and only non-zero b_{ik} entries on **B** produce non-zero entries on $\widetilde{\mathbf{B}}$.[5]

[3] Data can be found at:
http://www.dst.dk/en/Statistik/emner/nationalregnskab-og-offentlige-finanser/produktivitet-og-input-output/input-output-tabeller.
[4] A similar matrix relating consumption with imports by industry is also available.
[5] The resulting matrix $\widetilde{\mathbf{B}}$ is conveniently re-scaled in order to guarantee that all their columns sum up to one.

In order to evaluate the effect of specifying a matrix $\widetilde{\mathbf{B}}$ that deviate from the true **B**, we compare the values of the output generated by a constant vector **c** of private consumption ($\tilde{\mathbf{x}}^{\mathbf{c}}$) with that observed if the bridge matrix applied was **B** ($\mathbf{x}^{\mathbf{c}}$). In particular, we compute the percent relative deviation

$$100(\tilde{x}_i^c - x_i^c)/x_i^c \tag{6}$$

for each industry $i = 1, \dots, 68$ in the 2015 domestic Danish IO table. This comparison is made through 1,000 simulation draws. Tables 1, 2 and 3 summarize the results for different values of scalar δ.

<<Insert Table 1 about here>>
<<Insert Table 2 about here>>
<<Insert Table 3 about here>>

The columns in these tables show the mean percent deviation, as well as indicators of their variability as the variance and the range between the minimum and maximum percent deviations through the simulations. The figures in the tables suggest a negligible average effect of misrepresentations on **B**, which is not surprising given that the disturbance error is simulated to have a zero mean. More interesting, however, is the analysis of the indicators of variability. The results reported on these tables show how relative minor errors in the specification of the true bridge matrix **B** can produce relatively large deviations on the output generated by private consumption. More specifically, even if the scalar δ is set to the minimum value considered in the experiments ($\delta = 0.05$), the error on the output estimated by an input-output model can be larger than $\pm 7\%$ for the "Oil refinery" or "Cultural Activities and Arts" industries, or more than $\pm 10\%$ for industries like "Water collection and supply" or "Sports activities, amusement and recreation".

For larger values of scalar δ the deviations between the true ($\mathbf{x}^{\mathbf{c}}$) and the estimated output ($\tilde{\mathbf{x}}^{\mathbf{c}}$) become, not surprisingly, larger and significant errors are also detected for many other industries. In summary, our results suggest that correct or wrong specifications of the bridge matrix can play a very relevant role on the quantification by means of input-output models of private consumption impacts.

4. Can we estimate B? Applying entropy econometrics with an empirical illustration

Once the relevance of a correct specification of the bridge matrix has been assessed, this section explores the possibilities of estimating its cells in a situation where the underlying true matrix is not made available to the researcher. Mathematically speaking, this is a general problem of matrix balancing, where the researchers needs to reconcile data on consumer expenditure survey with data on an IO database (Steen-Olsen et al., 2016). In particular, we assume a situation where the vector of consumption (**c**) from a household survey and the vector of final private consumption by industry in the IO table ($\mathbf{y}^{\mathbf{c}}$) are both observable. Additionally, the researcher has some imperfect knowledge of the bridge matrix $\mathbf{B}$ that links both vectors.[6] This approximate matrix is denoted as $\widetilde{\mathbf{B}}$, similarly to the previous section.

For proposing an estimation technique that exploits all the available information, we mainly base on the paper by Golan et al. (1994). This paper presented a Cross Entropy (CE) procedure to estimate intersectoral flows from incomplete data or, more generally speaking, for a problem of matrix balancing with partial information. This technique bases on considering each column of the bridge matrix $\mathbf{B}$ as a probability distribution – note that their cells are positive and summing up to one- to be estimated. Similarly, the cells on $\widetilde{\mathbf{B}}$ are considered as an initial probability distribution. Similar to RAS, the CE technique minimizes the Kullback-Leibler divergence between the target $\mathbf{B}$ and the initial $\widetilde{\mathbf{B}}$, provided that the solution is consistent with the observable information –vectors $\mathbf{c}$ and $\mathbf{y^c}$ in our problem-. A constrained minimization problem is applied in order to find the solution to the CE estimator. The estimation problem can be posed as a minimization program like:

$$\underset{\boldsymbol{B}}{Min}\, D\left(\mathbf{B}, \widetilde{\mathbf{B}}\right) = \sum_{i=1}^{n}\sum_{k=1}^{p} b_{ik} ln\left(\frac{b_{ik}}{\tilde{b}_{ik}}\right) \tag{7}$$

Subject to:

$$\sum_{k=1}^{p} b_{ik}c_k = y_i^c;\ i = 1, \dots, n \tag{8}$$

$$\sum_{k=1}^{p} b_{ik} = 1;\ k = 1, \dots, p \tag{9}$$

The CE program depicted above is known to produce the same solution as a RAS adjustment of the initial $\widetilde{\mathbf{B}}$ matrix given vectors $\mathbf{c}$ and $\mathbf{y^c}$. However, while a RAS adjustment gives a "deterministic" solution, a CE solution makes possible doing some inference with the estimates, following Golan et al. (1994). Once the target bridge matrix is estimated, the CE framework makes possible hypothesis testing, basing on the relationship between the objective functions of restricted and unrestricted CE problems.

Let $D_U\left(\mathbf{B}, \widetilde{\mathbf{B}}\right) = \sum_{i=1}^{n}\sum_{k=1}^{p} b_{ik} ln\left(\frac{b_{ik}}{\tilde{b}_{ik}}\right)$ be the Kullback-Leibler divergence evaluated at the solution of the optimization problem as in equations (7) to (9) and $D_R\left(\mathbf{B}, \widetilde{\mathbf{B}}\right)$ be the same function where the solutions are restricted to fulfil J additional constraints –an example with $J = 1$ could be to test that a specific cell b_{ik} is equal to some arbitrary scalar-. Under some mild assumptions, (Golan et al., 2000, pp. 407–408) it is possible to obtain the following statistics:

$$2\left[D_R\left(\mathbf{B}, \widetilde{\mathbf{B}}\right) - D_U\left(\mathbf{B}, \widetilde{\mathbf{B}}\right)\right] \to \chi_J^2 \tag{10}$$

[6] This knowledge can be given by a past bridge matrix, by observing a bridge matrix in other economy that is expected to be similar to the economy of interest or by subjectively assigning values to its entries basing on some qualitative mapping.

To illustrate how the CE approach works, we use a real world case as illustration, making use again of the series of IO data published by Denmark Statistics. In particular, we will estimate the bridge matrix for the Danish economy on 2015 basing on the observable totals of household consumption by COICOP product ($\mathbf{c}$) and the vector of final private consumption reported on the IO tables ($\mathbf{y^c}$), both for 2015. The advantage of conducting this estimation exercise is that it allows for evaluating the accuracy of our estimates, since the true values of the target matrix **B** are actually observable.

Regarding the specification of the initial matrix $\widetilde{\mathbf{B}}$**,** in our analysis we have considered two alternative scenarios. First, we have assumed a situation with minimal information about the plausible values on the entries of the target matrix **B**, considering that the only available information to link household surveys with IO data is the qualitative correspondence between the COICOP and the CPA –or equivalent classification used on the IO tables for the economy of interest-. Note that this only allows for identifying those cells on **B** for which we expect to have non-zero values, but we do not have any information about the plausible proportions. In such a case, all the columns on $\widetilde{\mathbf{B}}$ will behave as uniform probability distributions for the non-zero entries. With such an initial $\widetilde{\mathbf{B}}$, the CE problem is equivalent to the Maximum Entropy (ME) solution that minimizes the Kullback-Leibler divergence with respect to an initial uniform distribution.

Alternatively, we have also considered a case with more information at hand in the form of a more informative initial matrix. In particular, we have studied the case where $\widetilde{\mathbf{B}}$ is set as the bridge matrix released by Denmark Statistics for 2010, which is expected to be similar, to the target matrix **B** for 2015.

With these data at hand, we have solved a minimization program as the one depicted in equations (7) to (9) in both scenarios, getting estimates for the cells of the matrix **B**, denoted as $\hat{b}_{ik}$. We have compared the ME and CE estimates with the true –but assumed as unobservable in the estimation exercise- b_{ik} in order to measure the accuracy of our estimation. In particular, we have calculated as a deviation measure the Weighted Absolute Percentage Error (WAPE) for each industry included on **B**, defined as:

$$WAPE_i = \sum_{k=1}^{p} 100 \frac{|b_{ik} - \hat{b}_{ik}|}{\sum_{k=1}^{p}|b_{ik}|}; i = 1, \dots, 68 \quad (11)$$

This measure has been largely used when evaluating non-survey input-output techniques (Temurshoev et al., 2011), since it averages the percentage error by weighting more the errors in larger cells (Oosterhaven et al., 2008). Table 4 shows the figures of this indicator by industry on each one of the two scenarios considered.

<<Insert Table 4 about here>>

The outcome of these empirical illustrations suggest that this estimation procedure, if the table taken as initial is expected to be similar to the target, can get relatively accurate estimates. The deviation figures where the initial $\widetilde{\mathbf{B}}$ matrix is similar to the true **B** are, generally speaking, considerably lower than in a situation where our initial assumptions about the structure of **B** are not informative. The results of our exercise highlight again the importance of having as much information as possible regarding the bridge matrices. One message would be that when conducting impact analysis of private consumption by IO models. If the statistical agencies do not publish series of official estimates of this matrix for the economy of interest, it is useful at least to have available other matrices – for previous periods or other economies- reasonably similar to the matrix **B** that can be used for estimation purposes.

5. Concluding remarks

Herrendorf et al. (2013) focus the attention on applying a consistent definition of commodities on both the household side when we estimate utility functions and the production side when we estimate production functions. Whereas in multisector models the sector classification is the same no matter which agent is using or producing them — i.e. the three broad sectors agriculture, manufacturing and services are the same for the consumption and production side—, the data show a different reality.

The proper connection may have significant effects not only on the analysis of structural change but also in the comparison across studies However, there are several issues that makes the construction of multisector models from data far to be an obvious task, being the concordance of both data structures a challenge for the researcher. The process to solve this challenge requires the construction of a concordance or bridge. However, this bridge matrix should be estimated since national bridge matrices are not available for all countries or only partial information is offered it the best cases.

Although Herrendorf et al. (2013) made an important and essential contribution to modern economic literature focused on structural change literature, the specific intricate input-output relationships between both data structures remains an unsolved challenge.

Our paper fills this gap by: first, providing an assessment of the importance of the correct specification of the bridge matrix; second, providing a technique that allows to obtain such bridge matrix when the necessary detailed data is not available.

Table 1. Percent deviation figures. 1,000 simulation draws. $\delta = 0.05$

Industry	max	mean	min	variance
Agriculture and horticulture	4.816	0.114	-4.338	3.412
Forestry	3.622	0.177	-3.001	1.673
Fishing	5.085	0.056	-4.284	3.240
Mining and quarrying	2.044	0.131	-2.042	0.510
Food products, bev.and tobacco	5.130	0.116	-4.586	4.144
Textiles and leather products	5.470	0.576	-5.313	4.428
Wood and wood products	1.078	0.137	-0.616	0.095
Manufacture of paper and paper products	1.930	0.547	-0.913	0.330
Printing etc.	0.913	0.182	-0.868	0.099
Oil refinery etc.	7.639	-0.126	-9.837	9.340
Manufacture of chemicals	2.482	0.309	-1.682	0.671
Pharmaceuticals	5.605	-0.122	-6.325	4.681
Manufacture of rubber and plastic products	1.388	0.287	-0.804	0.191
Other non-metallic mineral products	1.374	0.175	-0.999	0.190
Manufacture of basic metals	0.983	0.315	-0.580	0.075
Fabricated metal products	1.737	0.735	-0.164	0.127
Manufacture of electronic components	3.389	-0.548	-3.688	1.667
Electrical equipment	1.833	0.112	-1.245	0.345
Manufacture of machinery	0.900	0.215	-0.581	0.072
Motor vehicles and related parts	2.871	0.294	-1.618	0.506
Ships and other transport equipment	1.794	0.038	-1.858	0.588
Furniture and other manufacturing	6.295	0.419	-4.848	3.763
Repair and installation of equip.	1.501	0.126	-1.273	0.299
Electricity, gas, steam and a.c.	1.662	0.060	-1.644	0.310
Water collection and supply	10.296	0.013	-9.315	13.925
Sewerage; waste collection, etc.	2.349	0.116	-2.679	0.919
Construction	0.546	0.125	-0.364	0.023
Trade and repair of motor vehicles	5.107	0.402	-3.759	2.085
Wholesale	1.786	-0.189	-1.756	0.352
Retail sale	3.561	0.892	-2.215	0.983
Land transport and pipelines	1.288	0.060	-0.908	0.164
Water transport	3.171	0.149	-2.239	0.869
Air transport	1.393	0.140	-1.246	0.337
Support activities for transportation	2.279	0.017	-2.260	0.626
Postal and courier activities	0.552	0.234	-0.172	0.019
Accommodation and food services	0.785	0.023	-0.885	0.081
Publishing activities	1.255	0.147	-1.445	0.227
Motion picture, tv., sound recording and radio	3.419	0.200	-3.725	1.633
Telecommunications	1.962	0.193	-1.517	0.543
IT and information service activities	0.668	0.229	-0.201	0.024
Financial services	0.314	0.066	-0.177	0.011
Insurance and pension funding	0.948	0.046	-1.506	0.179

Table 1 (continued). Percent deviation figures. 1,000 simulation draws. $\delta = 0.05$

Industry	max	mean	min	variance
Other financial activities	1.561	0.078	-0.969	0.172
Buying and selling of real estate	4.749	0.071	-4.441	2.032
Renting of non-residential buildings	0.956	0.370	-0.326	0.050
Renting of residential buildings	0.172	0.000	-0.205	0.004
Owner-occupied dwellings	0.128	0.000	-0.104	0.002
Legal and accounting activities	0.723	0.345	-0.097	0.014
Architectural and engineering activities	0.680	0.180	-0.469	0.041
Scientific research and development (market)	2.228	0.132	-1.302	0.473
Scientific research and development (non-market)	0.590	0.212	-0.307	0.024
Advertising and market research	0.652	0.256	-0.262	0.033
Other professional, scientific activities	2.288	0.708	-0.603	0.275
Rental and leasing activities	2.717	1.028	-0.508	0.264
Employment activities	0.680	0.314	-0.093	0.020
Travel agent activities	0.281	0.016	-0.282	0.011
Other business service activities	1.541	0.873	0.075	0.060
Public administration ect.	2.012	0.349	-1.947	0.343
Rescue service ect. (market)	6.977	-0.096	-8.031	8.050
Education (non-market)	1.589	0.054	-1.667	0.341
Adult and other education (market)	8.842	0.179	-9.156	8.789
Human health activities	0.409	0.011	-0.372	0.022
Residential care	2.535	0.081	-1.660	0.522
Cultural activities, Arts, etc.	7.201	-0.222	-7.595	8.406
Sports activities, amusement and recreation	11.629	0.215	-10.547	13.465
Activities of membership organizations	3.883	0.238	-2.904	1.236
Repair of personal goods	3.698	-0.217	-4.592	2.371
Other personal service activities	4.867	0.112	-3.152	2.746

Table 2. Percent deviation figures. 1,000 simulation draws. $\delta = 0.10$

Industry	max	mean	min	variance
Agriculture and horticulture	9.732	0.202	-8.536	13.737
Forestry	7.354	0.323	-6.070	6.767
Fishing	10.286	0.095	-8.427	13.054
Mining and quarrying	3.791	0.061	-4.214	2.032
Food products, bev.and tobacco	10.375	0.200	-9.020	16.687
Textiles and leather products	10.940	0.668	-10.643	17.842
Wood and wood products	2.006	0.146	-1.317	0.384
Manufacture of paper and paper products	3.441	0.601	-2.264	1.328
Printing etc.	1.588	0.172	-2.011	0.403
Oil refinery etc.	15.589	-0.282	-19.419	37.360
Manufacture of chemicals	4.833	0.388	-3.714	2.710
Pharmaceuticals	11.529	-0.228	-12.555	18.802
Manufacture of rubber and plastic products	2.556	0.298	-1.914	0.769
Other non-metallic mineral products	2.694	0.239	-1.972	0.771
Manufacture of basic metals	1.687	0.321	-1.438	0.302
Fabricated metal products	2.785	0.747	-1.016	0.510
Manufacture of electronic components	7.537	-0.577	-6.897	6.699
Electrical equipment	3.584	0.104	-2.663	1.370
Manufacture of machinery	1.571	0.211	-1.347	0.289
Motor vehicles and related parts	5.445	0.315	-3.703	2.027
Ships and other transport equipment	3.658	0.162	-3.609	2.378
Furniture and other manufacturing	12.287	0.428	-9.993	15.041
Repair and installation of equip.	2.958	0.150	-2.597	1.203
Electricity, gas, steam and a.c.	3.087	0.005	-3.782	1.265
Water collection and supply	22.319	0.119	-18.447	56.219
Sewerage; waste collection, etc.	4.427	0.099	-5.947	3.728
Construction	0.912	0.110	-0.879	0.093
Trade and repair of motor vehicles	9.754	0.409	-8.319	8.366
Wholesale	3.765	-0.132	-3.217	1.410
Retail sale	6.002	0.823	-5.609	3.936
Land transport and pipelines	2.445	0.054	-1.930	0.667
Water transport	6.415	0.225	-4.259	3.562
Air transport	2.854	0.187	-2.514	1.368
Support activities for transportation	4.622	0.045	-4.509	2.503
Postal and courier activities	0.871	0.242	-0.579	0.079
Accommodation and food services	1.417	-0.026	-2.158	0.345
Publishing activities	2.257	0.119	-3.192	0.926
Motion picture, tv., sound recording and radio	6.849	0.257	-7.516	6.528
Telecommunications	3.824	0.258	-3.270	2.177
IT and information service activities	1.140	0.254	-0.600	0.095
Financial services	0.575	0.068	-0.414	0.043
Insurance and pension funding	1.847	0.037	-3.580	0.757

Table 2 (continued). Percent deviation figures. 1,000 simulation draws. $\boldsymbol{\delta = 0.10}$

Industry	**max**	**mean**	**min**	**variance**
Other financial activities	3.569	0.096	-1.878	0.726
Buying and selling of real estate	9.887	0.108	-8.75	8.158
Renting of non-residential buildings	1.497	0.354	-1.1	0.202
Renting of residential buildings	0.376	0.004	-0.375	0.016
Owner-occupied dwellings	0.234	-0.003	-0.227	0.006
Legal and accounting activities	1.125	0.352	-0.526	0.056
Architectural and engineering activities	1.193	0.185	-1.123	0.165
Scientific research and development (market)	4.483	0.126	-2.71	1.903
Scientific research and development (non-market)	1.008	0.228	-0.761	0.097
Advertising and market research	1.07	0.264	-0.766	0.133
Other professional, scientific activities	3.98	0.775	-1.837	1.106
Rental and leasing activities	4.443	1.018	-2.022	1.053
Employment activities	1.092	0.337	-0.47	0.081
Travel agent activities	0.567	0.025	-0.554	0.044
Other business service activities	2.308	0.91	-0.668	0.243
Public administration ect.	3.739	0.437	-3.894	1.408
Rescue service ect. (market)	14.004	-0.167	-16.043	32.149
Education (non-market)	2.778	-0.035	-3.942	1.386
Adult and other education (market)	19.114	0.472	-17.86	35.312
Human health activities	0.731	-5.70E-04	-0.826	0.089
Residential care	5.836	0.214	-3.077	2.222
Cultural activities, Arts, etc.	13.615	-0.531	-15.415	33.729
Sports activities, amusement and recreation	23.334	0.417	-21.482	53.951
Activities of membership organizations	7.504	0.243	-6.158	4.99
Repair of personal goods	8.069	-0.236	-9.046	9.483
Other personal service activities	9.297	0.133	-6.813	11.149

Table 3. Percent deviation figures. 1,000 simulation draws. $\delta = 0.15$

Industry	max	mean	min	variance
Agriculture and horticulture	14.788	0.292	-12.572	27.36
Forestry	11.294	0.518	-9.164	20.458
Fishing	15.638	0.141	-12.409	28.047
Mining and quarrying	5.439	-0.029	-6.335	11.774
Food products, bev.and tobacco	15.771	0.279	-13.947	29.718
Textiles and leather products	17.728	0.901	-15.449	33.177
Wood and wood products	2.928	0.166	-1.971	4.899
Manufacture of paper and paper products	5.034	0.661	-3.66	8.693
Printing etc.	2.22	0.155	-3.266	5.485
Oil refinery etc.	24.161	-0.415	-28.702	52.864
Manufacture of chemicals	7.32	0.493	-5.892	13.212
Pharmaceuticals	17.86	-0.269	-18.652	36.512
Manufacture of rubber and plastic products	3.794	0.323	-3.044	6.838
Other non-metallic mineral products	4.092	0.316	-2.821	6.913
Manufacture of basic metals	2.429	0.329	-2.257	4.687
Fabricated metal products	3.871	0.763	-1.832	5.703
Manufacture of electronic components	11.99	-0.567	-10.108	22.098
Electrical equipment	5.404	0.125	-4.111	9.515
Manufacture of machinery	2.242	0.213	-2.07	4.312
Motor vehicles and related parts	7.986	0.32	-6.032	14.017
Ships and other transport equipment	5.839	0.34	-5.306	11.146
Furniture and other manufacturing	18.486	0.522	-14.953	33.438
Repair and installation of equip.	4.475	0.173	-3.875	8.35
Electricity, gas, steam and a.c.	4.387	-0.089	-6.528	10.915
Water collection and supply	36.601	0.355	-27.373	63.974
Sewerage; waste collection, etc.	6.367	0.06	-9.843	16.211
Construction	1.241	0.088	-1.411	2.652
Trade and repair of motor vehicles	14.31	0.376	-13.396	27.706
Wholesale	5.724	-0.047	-4.61	10.334
Retail sale	8.265	0.715	-9.302	17.567
Land transport and pipelines	3.532	0.03	-3.026	6.558
Water transport	9.874	0.353	-6.007	15.881
Air transport	4.563	0.276	-3.683	8.247
Support activities for transportation	7.071	0.096	-6.735	13.806
Postal and courier activities	1.193	0.252	-0.997	2.19
Accommodation and food services	1.966	-0.101	-3.997	5.963
Publishing activities	3.18	0.076	-5.115	8.296
Motion picture, tv., sound recording and radio	10.46	0.329	-11.219	21.679
Telecommunications	5.716	0.327	-5.307	11.023
IT and information service activities	1.625	0.286	-0.989	2.613
Financial services	0.843	0.067	-0.651	1.494
Insurance and pension funding	2.721	-0.009	-6.555	9.275

Table 3 (continued). Percent deviation figures. 1,000 simulation draws. $\boldsymbol{\delta = 0.15}$

Industry	**max**	**mean**	**min**	**variance**
Other financial activities	6.476	0.156	-2.662	9.138
Buying and selling of real estate	15.545	0.175	-12.878	28.423
Renting of non-residential buildings	2.004	0.327	-1.964	3.968
Renting of residential buildings	0.619	0.014	-0.519	1.137
Owner-occupied dwellings	0.323	-0.009	-0.375	0.698
Legal and accounting activities	1.55	0.36	-0.949	2.5
Architectural and engineering activities	1.71	0.189	-1.791	3.502
Scientific research and development (market)	6.954	0.136	-4.069	11.023
Scientific research and development (non-market)	1.484	0.252	-1.181	2.665
Advertising and market research	1.5	0.27	-1.27	2.77
Other professional, scientific activities	5.736	0.854	-3.052	8.788
Rental and leasing activities	6.234	1.02	-3.496	9.73
Employment activities	1.54	0.367	-0.839	2.379
Travel agent activities	0.869	0.039	-0.806	1.675
Other business service activities	3.165	0.957	-1.386	4.551
Public administration ect.	5.881	0.577	-5.626	11.507
Rescue service ect. (market)	21.149	-0.147	-23.985	45.134
Education (non-market)	3.755	-0.16	-6.965	10.72
Adult and other education (market)	31.605	0.91	-26.132	57.737
Human health activities	0.999	-0.022	-1.455	2.455
Residential care	10.313	0.409	-4.303	14.615
Cultural activities, Arts, etc.	19.379	-0.908	-24.51	43.889
Sports activities, amusement and recreation	35.162	0.651	-32.777	67.94
Activities of membership organizations	11.077	0.226	-9.562	20.639
Repair of personal goods	13.14	-0.165	-13.482	26.622
Other personal service activities	13.901	0.123	-11.024	24.925

Table 4. WAPE (%) for the 68 Danish Industries.

Industry	$WAPE_i$ (ME solution)	$WAPE_i$ (CE solution)
Agriculture and horticulture	3.073	1.193
Forestry	0.236	0.208
Fishing	3.485	1.816
Mining and quarrying	8.887	9.874
Food products, bev.and tobacco	3.601	4.064
Textiles and leather products	5.404	1.165
Wood and wood products	6.065	1.625
Manufacture of paper and paper products	9.338	8.273
Printing etc.	15.604	2.287
Oil refinery etc.	3.800	0.801
Manufacture of chemicals	8.799	1.378
Pharmaceuticals	1.861	0.425
Manufacture of rubber and plastic products	3.473	4.617
Other non-metallic mineral products	4.868	0.641
Manufacture of basic metals	12.595	0.746
Fabricated metal products	5.014	0.661
Manufacture of electronic components	7.519	0.228
Electrical equipment	4.128	1.226
Manufacture of machinery	5.523	1.221
Motor vehicles and related parts	2.629	4.189
Ships and other transport equipment	1.922	3.073
Furniture and other manufacturing	4.382	4.689
Repair and installation of equip.	33.772	39.840
Electricity, gas, steam and a.c.	2.063	0.002
Water collection and supply	0.273	0.001
Sewerage; waste collection, etc.	0.237	0.018
Construction	1.719	0.350
Trade and repair of motor vehicles	2.932	0.180
Wholesale	0.994	0.615
Retail sale	0.885	0.368
Land transport and pipelines	0.613	0.010
Water transport	0.418	0.056
Air transport	0.813	0.320
Support activities for transportation	8.916	0.830
Postal and courier activities	0.008	0.006
Accommodation and food services	0.009	0.002
Publishing activities	2.044	0.115
Motion picture, tv., sound recording and radio	0.716	0.122
Telecommunications	0.833	0.016
IT and information service activities	4.865	1.151
Financial services	0.089	0.002
Insurance and pension funding	0.000	0.000

Table 4 (continued). WAPE (%) for the 68 Danish Industries.

Industry	$\boldsymbol{WAPE_i}$ **(ME solution)**	$\boldsymbol{WAPE_i}$ **(CE solution)**
Other financial activities	0.871	0.023
Buying and selling of real estate	0.365	0.059
Renting of non-residential buildings	8.989	0.049
Renting of residential buildings	0.007	0.001
Owner-occupied dwellings	0.000	0.000
Legal and accounting activities	3.949	1.905
Architectural and engineering activities	3.758	0.033
Scientific research and development (market)	4.750	0.423
Scientific research and development (non-market)	0.000	0.000
Advertising and market research	8.834	0.227
Other professional, scientific activities	1.513	0.551
Rental and leasing activities	1.360	0.549
Employment activities	1.556	0.174
Travel agent activities	0.049	0.004
Other business service activities	2.101	0.105
Public administration ect.	1.499	0.134
Rescue service ect. (market)	6.674	0.074
Education (non-market)	0.000	0.000
Adult and other education (market)	1.337	0.006
Human health activities	0.267	0.014
Residential care	0.000	0.000
Cultural activities, Arts, etc.	2.449	0.012
Sports activities, amusement and recreation	1.030	0.356
Activities of membership organizations	0.817	0.038
Repair of personal goods	1.115	0.540
Other personal service activities	1.540	0.061

References

Baumol, W.J., 1967. "Macroeconomics of unbalanced growth: the analonomy of urban crisis". Am. Econ. Rev. 57(3), 415-426.

Golan, A., Judge, G., Robinson, S., 1994. "Recovering information from incomplete or partial multisectoral economic data". Rev. Econ. Stat. 76, 541-549.

Golan, A., Karp, L.S., Perloff, J.M., 2000. "Estimating Coke’s and Pepsi’s Price and Advertising Strategies". J. Bus. 18, 398–409. doi:10.2307/1392222Kuznets, S., 1966. *Modern Economic Growth*. New Haven: Yale University Pres.

Herrendorf, B., Herrington, C., Valentinyi, A., 2015. "Sectoral technology and structural transformation". Am. Econ. J-Macroecon. 7(4), 104-133.

Herrendorf, B., Rogerson, R., Valentinyi, A., 2014. "Growth and structural transformation". In *Handbook of economic growth*, vol. 2. Ed. P. Aghion, S. Durlauf. Amsterdam [etc.]: Elsevier.

Herrendorf, B., Rogerson, R., Valentinyi, A., 2013. "Two perspectives on preferences and structural transformation". Am. Econ. Rev. 103(7), 2752-2789.

Kongsamut, P., Rebelo, S., Xie, D., 2001. "Beyond balanced growth". Rev. Econ. Stud. 68(4), 869-882.

Kuznets, S., 1966. *Modern economic growth*. New Haven: Yale University Press.

Kuznets, S., 1973. "Modern Economic Growth: Findings and Reflections". Am. Econ. Rev. 63, 247-258.

Ngai, L.R., Pissarides, C.A., 2007. "Structural change in a multisector model of growth". Am. Econ. Rev. 97(1), 429-443.

Oosterhaven, J., Stelder, D., Inomata, S., 2008. "Estimating international interindustry linkages: Non-survey simulations of the Asian-pacific economy". Econ. Syst. Res. 20, 395–414. doi:10.1080/09535310802551448

Steen-Olsen, K., Wood, R., Hertwich, E.G., 2016. "The Carbon Footprint of Norwegian Household Consumption 1999-2012". J. Ind. Ecol. 20, 582–592. doi:10.1111/jiec.12405

Temurshoev, U., Webb, C., Yamano, N., 2011. "Projection of Supply and Use Tables: Methods and Their Empirical Assessment". Econ. Syst. Res. 23, 91–123. doi:10.1080/09535314.2010.534978

---

[i] Herrendorf et al. (2014) provide a review of recent works that apply multisector models to analyse structural change.

[ii] When we focus on products rather than on industries the proper classification is the *Central product classification* (CPC) that is closely related with the ISIC.

[iii] https://www.bea.gov/scb/pdf/2008/05%20May/0508_nipa_pce.pdf

[iv] https://www.census.gov/eos/www/naics/