UNU-MERIT Working Paper Series

#2014-070

Analysing global value chains using input-output economics: Proceed with care

Önder Nomaler and Bart Verspagen

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)
email: info@merit.unu.edu | website: http://www.merit.unu.edu

Maastricht Graduate School of Governance (MGSoG)
email: info-governance@maastrichtuniversity.nl | website: http://mgsog.merit.unu.edu

Keizer Karelplein 19, 6211 TC Maastricht, The Netherlands
Tel: (31) (43) 388 4400, Fax: (31) (43) 388 4499

UNU-MERIT Working Papers

ISSN 1871-9872

Maastricht Economic and social Research Institute on Innovation and Technology, UNU-MERIT

Maastricht Graduate School of Governance MGSoG

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT and MGSoG to stimulate discussion on the issues raised.

Analysing Global Value Chains Using Input-Output Economics: Proceed With Care

Önder Nomaler*,** and Bart Verspagen*,***

* UNU-MERIT, Maastricht

** Eindhoven University of Technology

*** Maastricht University, Department of Economics

Abstract

Input-output economics has become a popular tool to analyse the international fragmentation of value chains, especially now that several multi-regional tables that cover large parts of the global economy have become available. It has been argued that these tables, when analysed with the help of the input-output economics toolbox, can provide better insights about global value chains than can be obtained by case studies of individual value chains. We argue that there are several problems related to the aggregated nature of the input-output table that may lead to large distortions and biases in the aggregate picture about global value chains that is obtained by input-output analysis. There are three main sources behind the distortion obtained in static decompositions of value chains: the average nature of value added to output ratios in the tables, the emergence of production cycles in the process of aggregating several value chains into a single table, and the characteristic of the so-called inverse Leontief matrix to even out the value added distribution. We provide an overview of how these distortions work, and argue that under a wide range of circumstances, input-output methods tend to overstate the contribution of the final sector to the value chain. We also show that this bias does not vanish when we compare input-output decompositions at two different points in time.

Keywords: Input-output economics; Global Value Chains; International fragmentation of production

JEL codes: C67, F62

3 October 2014

We thank Bart Los and Adam Szirmai for helpful discussions while preparing this draft. The views expressed, and any potential errors therein, can only be attributed to the authors.

1. Introduction

Today's manufactured products are often produced in a globally dispersed way, with increasingly finished products and parts being shipped between different parts of the globe. Firms' decisions on where to locate production are based on profit considerations. Capabilities of the firm itself and those of potential suppliers elsewhere, as well as differences in factor costs and transport costs between locations, are factors that drive decisions on where to locate parts of these global value chains.

Empirical evidence suggests that the outcome of this process is that different locations capture largely varying parts of the value created in individual products (e.g., Shin et al., 2010). Data on some individual but well-known cases, such as the iPod or mobile phones, suggest that the largest part of value added may be created in activities such as R&D and marketing, rather than in actual production (e.g., Ali-Yrrkö et al, 2011; Dedrick et al., 2099). This implies that a rapid rise in manufacturing capability may not automatically translate into rapid productivity growth, or rapid growth of living standards. Thus, if Asia (China) is turning into the factory of the world, this may not have strong implications for living standards.

But the case studies of global value chains for individual products are not necessarily representative for the broad manufacturing base that is being developed in China or other countries. It may well be the case that manufacturers capture more value in other, less prolific value chains, or that the newly emerging economies participate in the parts of other global value chains that yield high value added. Timmer et al. (2014) suggest that this is indeed the case. They base their conclusions on input-output analysis and a new data source, the World Input-Output Database (WIOD). Because the WIOD captures the entire economy of the 40 countries it covers, it may be expected to give an overall, and thus more balanced picture, or what global value chains imply for emerging economies such as China, or developed economies such as the US or Europe.

Timmer et al. (2014) and Los et al. (2014), or input-output analysis in general, try to decompose the value added created in response to aggregate demand, into country-sector combinations. For example, whereas a case study of the global value chain for iPods would give an overview of which producers in which locations capture how much value from selling one device, Timmer et al. (2014) make an attempt to calculate how the 1$ value that is associated to the sale of 1$ worth of products from the entire Chinese electronics industry is distributed among 35 sectors located in 40 countries. The difference lies in the level of aggregation: whereas the case study looks at a single product only, the input-output analysis is supposed to give an adequate picture of the entire range of electronics products (or any different industrial sector) produced in a specific location.

In this paper, we argue that the aggregate picture provided by input-output analysis may be significantly distorted. A central tenet of our argument is that input-output analysis will, under a wide range of circumstances, over-estimate the contribution of value added made by the sector that produces the final product, i.e., the one closest to final demand. Thus, rather than providing a correct representation of the average value added distribution of various global value chains, the decompositions of value based on input-output economics may significantly misrepresent the global distribution of value added generation.

However, it is not only the value added contribution of the last sector in the value chain that is distorted. The source of distortion, which we will illustrate and discuss in detail, applies to all sectors that are involved in production. Hence any statistic that looks at the broad distribution of value added created in response to a specific category of demand, will have some degree of distortion. This holds in particular for the so-called Foreign Value Added Share (FVAS) of Timmer et al. (2014) and Los et al. (2014). One minus FVAS measures the share of all domestic sectors to value added created by (final) demand served by a sector. This is similar to the measures used in some of the case studies, such as the analysis on mobile phones by Ali-Yrrkö et al. (2011).

The root of our argument is concerned with aggregation errors, which is not an unknown issue in input-output economics (e.g., Miller and Blair, 2009, chapter 4). However, in the context of global value chains and fragmented production systems, the aggregation issue takes special importance, and is of a somewhat different nature than what has usually been debated in the literature. With increasing (international) fragmentation, it is likely that a single sector contributes in various stages of the production process. If a sector contributes 1$ value added to a value chain, the share of total value added by the sector to the entire chain should not be affected by whether the 1$ is added at the beginning of the chain (downstream) or at the end of the chain (upstream). However, in terms of the value-added-to-output ratio of the sector, this matters a great deal. The 1$ represents a high (low) value added to output ratio if it is added at the beginning (end) of the chain. Input output analysis assumes, however, that this ratio is fixed for all activities (value chains) that the sector is involved in. We show that the heavy dependence of input-output analysis on this assumption creates an aggregation error that is far from “neutral” white noise that leaves the average of a “representative” value chain unaffected.

Our analytical exposition is aimed at very simple cases (3 sectors, short value chains without branching), and only covers a static perspective (it looks at decompositions of aggregated value chains in a single period). However, the simulations that we undertake after the analytical exposition also cover longer value chains in a setting with more sectors, and provide a dynamic perspective by looking at potential distortion that arises in comparing 2 periods. For the static perspective, the simulations confirm the conclusions from the analytical work, i.e., that the input-output decompositions provide a distorted picture, often over-estimating the value added contribution of the final sector, but with a significant distortion for all sectors that contribute to the chain. The static simulations also confirm that the FVAS statistic is significantly distorted.

The dynamic simulations show that the distortion also carries over to comparisons of multiple periods. Here we focus exclusively on the change in the FVAS statistic and show that the input-output decompositions tend to particularly over-estimate the increase in FVAS of the final sectors where FVAS was initially low. In this form of distortion, an additional source of bias reveals itself. This is the tendency of the inverse Leontief matrix, which is used in the input-output calculations, to even out the value added distribution. An extreme form of this phenomenon, which we refer to as regression to the mean, is the fact that the input-output decompositions tend to identify a value added contribution from all sectors in the system to any specific value chain, while in the true value added distribution, only a few sectors contribute to any given specific value chain.

We proceed, in Section 2, to analytically explain some of the problems with value chain analysis by input-output analysis. To this end, we employ a number of simple examples of value chains, which we aggregate and then analyse using the standard input-output methods. The results will show how these calculations provide a distorted picture of the underlying value chains that we aggregated. We also provide an analytical exploration of the nature and extent of the distortion, and find that there are two main sources. In Section 3, we provide simulation experiments, in which we aggregate many more value chains. The aim of this section is to see whether the problems that we point to in Section 2 also exist in example value chains that are more complex and more numerous than the simple examples we present in Section 3. In the concluding Section 4, we provide a brief discussion of the results, and call for more empirical evidence at the level of value chains to assess how large the distortions that we point to really are.

2. The problem

The aggregated nature of input-output tables confronts serious problems for the analysis of value chains, due to two main conceptual issues. The first of these is the sequential nature of value added in a value chain. This leads to a steady increase in gross value (of output) when a product progresses in the chain, which makes it problematic to use gross output (of a firm or a sector that contributes to the chain) as a benchmark to scale the value added contribution of the firm of sector. Using (average) value added-to-output ratios is problematic in comparing a firm that adds one unit of value added at the beginning of the value chain to a firm that adds one unit of value added at the end of the chain.

The second problem relates to the notion of production cycles in the value chain. We define a production cycle as a case where a product is produced using the product itself, either directly or indirectly. A nuclear power plant is an example of a case where production cycles exist: it uses electricity to produce electricity. Other examples are harder to produce, if we limit ourselves to the “micro” level, i.e., the level of actual inputs and outputs at the firm level. At that level, many of the production cycles that exist in an aggregate input-output table turn out to be the result of aggregation. Take, for example, a value chain in which wood is delivered to the fabricated metal sector as a packing material, whereas the fabricated metal sector delivers screws (packaged in plastic) to the wood & products sector, where they are used to produce furniture. There are no production cycles within each of the chains, but when we aggregate the chains to sectors, a cycle has emerged.

We proceed by analysing a few simple cases for value chains with 3 stages. We first derive some analytical conclusions for these chains, and then discuss some numerical examples.

2.1. Analysing some simple cases

Assume that in one value chain, sector A delivers total value Z_1 to sector B, which uses it to deliver total value Z_2 to sector C, and finally sector C uses the input it received from B to deliver Z_3 to final demand. We follow standard input-output logic to analyse this value chain. We denote the matrix of intermediate deliveries by **U**, the vector of value added by **V**, the vector of final demand by **F**, and the vector of gross output by **Q**. Elements of vectors or matrices are denoted by small case letters. For the above-described value chain, we have:

$$\mathbf{U} = \begin{bmatrix} 0 & Z_1 & 0 \\ 0 & 0 & Z_2 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 \\ 0 \\ Z_3 \end{bmatrix}, \mathbf{Q} = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix}$$

The matrix of input coefficients is **A**, where $a_{ij} = u_{ij}/q_j$, and **M** is a diagonal matrix with elements $m_i = v_i/q_i$ on the diagonal and zeros otherwise.

$$\mathbf{A} = \begin{bmatrix} 0 & \frac{Z_1}{Z_2} & 0 \\ 0 & 0 & \frac{Z_2}{Z_3} \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-\frac{Z_1}{Z_2} & 0 \\ 0 & 0 & 1-\frac{Z_2}{Z_3} \end{bmatrix}$$

The inverse Leontief matrix is defined in the conventional way, $\mathbf{L} = [\mathbf{I}-\mathbf{A}]^{-1}$:

$$\mathbf{L} = \begin{bmatrix} 1 & \frac{Z_1}{Z_2} & \frac{Z_1}{Z_3} \\ 0 & 1 & \frac{Z_2}{Z_3} \\ 0 & 0 & 1 \end{bmatrix}$$

The input-output way of looking at the value chain is to construct a matrix $\mathbf{D} = \mathbf{M}\mathrm{x}\mathbf{L}$, which decomposes value added: d_{ij} is the share of sector i in the value chain that produces a product from sector j. Note that the columns of matrix **D** will add to 1 (or zero, in case they do not contain any non-zero values), which is the reflection of the fact that the sum of final demand in the table is equal to the sum of value added.

Finally, let us introduce the diagonal matrix **Φ**, which has vector **F** on the main diagonal and zeroes otherwise, and also define the matrix **S** = **M**x**L**x**Φ = D**x**Φ.** Column *j* of **S** also provides a decomposition of value added in the chain, but in terms of dollars instead of shares. In other words, while each column of **D** adds up to 1, column *j* of **S** adds up to f_j which is the actual final demand for sector *j*. Similarly, each row *i* of **S** adds up to v_i, the total value added produced in sector *i*. For the case under analysis, we have the rather straightforward result:

$$\mathbf{D} = \begin{array}{|c|c|c|} \hline 1 & \frac{Z_1}{Z_2} & \frac{Z_1}{Z_3} \\ \hline 0 & 1-\frac{Z_1}{Z_2} & \frac{Z_2-Z_1}{Z_3} \\ \hline 0 & 0 & 1-\frac{Z_2}{Z_3} \\ \hline \end{array}, \quad \mathbf{S} = \begin{array}{|c|c|c|} \hline 0 & 0 & Z_1 \\ \hline 0 & 0 & Z_2 - Z_1 \\ \hline 0 & 0 & Z_3 - Z_2 \\ \hline \end{array}$$

Let us now introduce a second value chain, one where the final good is produced in sector B, whereas sector C provides the intermediate stage. This is a chain A → C → B, whereas the case before was a chain A → B → C. We think of this second value chain as one that uses and produces different goods than the ones in the first chain, but at the level of the input-output table, these goods would be aggregated into the same sectors as before. For example, while the final good in the first chain could be an iPod, the intermediate good that sector C supplies in the second chain could be a memory chip.

To keep things analytically traceable, we start by assuming that the values of gross output in this chain are exactly the same, at every stage, as in the first chain.[1] The second chain can be analysed like the first one, and it yields

$$\mathbf{D^*} = \begin{array}{|c|c|c|} \hline 1 & \frac{Z_1}{Z_3} & \frac{Z_1}{Z_2} \\ \hline 0 & 1-\frac{Z_2}{Z_3} & 0 \\ \hline 0 & \frac{Z_2-Z_1}{Z_3} & 1-\frac{Z_1}{Z_2} \\ \hline \end{array}, \quad \mathbf{S^*} = \begin{array}{|c|c|c|} \hline 0 & Z_1 & 0 \\ \hline 0 & Z_3 - Z_2 & 0 \\ \hline 0 & Z_2 - Z_1 & 0 \\ \hline \end{array}$$

Now we aggregate the two value chains in one input-output table, and apply the same decompositions, with the aim of finding out how much value is produced by each sector, in each of the two underlying chains. For the aggregated table, we have:

$$\mathbf{A^{agg}} = \begin{array}{|c|c|c|} \hline 0 & \frac{Z_1}{Z_2+Z_3} & \frac{Z_1}{Z_2+Z_3} \\ \hline \end{array}, \quad \mathbf{M^{agg}} = \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline \end{array}$$

[1] We will relax this assumption below when we apply simulation analysis, and also in one analytical case that we consider next.

0	0	$\frac{Z_2}{Z_2+Z_3}$
0	$\frac{Z_2}{Z_2+Z_3}$	0

0	$1-\frac{Z_1+Z_2}{Z_2+Z_3}$	0
0	0	$1-\frac{Z_1+Z_2}{Z_2+Z_3}$

$\mathbf{L}^{\mathbf{agg}}$ =

1	$\frac{Z_1}{Z_3}$	$\frac{Z_1}{Z_3}$
0	$\frac{(Z_2+Z_3)^2}{2Z_2Z_3+Z_3^2}$	$\frac{Z_2(Z_2+Z_3)}{2Z_2Z_3+Z_3^2}$
0	$\frac{Z_2(Z_2+Z_3)}{2Z_2Z_3+Z_3^2}$	$\frac{(Z_2+Z_3)^2}{2Z_2Z_3+Z_3^2}$

$\mathbf{D}^{\mathbf{agg}}$ =

1	$\frac{Z_1}{Z_3}$	$\frac{Z_1}{Z_3}$
0	$\frac{(Z_3-Z_1)(Z_2+Z_3)}{2Z_2Z_3+Z_3^2}$	$\frac{Z_2(Z_3-Z_1)}{2Z_2Z_3+Z_3^2}$
0	$\frac{Z_2(Z_3-Z_1)}{2Z_2Z_3+Z_3^2}$	$\frac{(Z_3-Z_1)(Z_2+Z_3)}{2Z_2Z_3+Z_3^2}$

$\mathbf{S}^{\mathbf{agg}}$ =

0	Z_1	Z_1
0	$\frac{(Z_3-Z_1)(Z_2+Z_3)}{2Z_2+Z_3}$	$\frac{Z_2(Z_3-Z_1)}{2Z_2+Z_3}$
0	$\frac{Z_2(Z_3-Z_1)}{2Z_2+Z_3}$	$\frac{(Z_3-Z_1)(Z_2+Z_3)}{2Z_2+Z_3}$

With all these results in place, we can ask what the aggregate decomposition tells us about the first value chain (A → B → C). We know that the true value added $ value (in matrix **S**) that is generated by Sector C in this chain is equal to Z_3- Z_2, but it is also seen that matrix $\mathbf{S}^{\mathbf{agg}}$ provides a different estimate of this value contribution: $\frac{(Z_3-Z_1)(Z_2+Z_3)}{2Z_2+Z_3}$. Similarly, the true share of value added (in matrix **D**) in this chain provided by Sector C is $1-\frac{Z_2}{Z_3}$, while matrix $\mathbf{D}^{\mathbf{agg}}$ puts the share at $\frac{(Z_3-Z_1)(Z_2+Z_3)}{2Z_2Z_3+Z_3^2}$.

We can derive the condition under which the aggregate input-output table overestimates these contributions by solving $Z_3 - Z_2 < \frac{(Z_3-Z_1)(Z_2+Z_3)}{2Z_2+Z_3}$ (obviously, this is the condition for the $ values, but the condition for the value added shares yields identical results). This yields $Z_3 < \frac{2Z_2^2}{Z_1} - Z_2$. If we rewrite $Z_2 = Z_1 + \Delta_2$ and $Z_3 = Z_1 + \Delta_2 + \Delta_3$, this condition can also be written as

$\Delta_3 < 2\left(\frac{\Delta_2^2}{Z_1} + \Delta_2\right)$, and normalising the $ value added contribution of the primary sector A to one, this simplifies further to $\Delta_3 < 2(\Delta_2^2 + \Delta_2)$. Thus, unless the final sector adds a rather large amount of value added to the chain, the aggregate input-output decomposition will overestimate the contribution of the final sector to total value added in the chain. Alternatively, if the final sector does contribute a really high amount of value added, the aggregate decomposition will underestimate its contribution. Only in the "knife edge" case when $\Delta_3 = 2\left(\frac{\Delta_2^2}{Z_1} + \Delta_2\right)$ will the aggregate decomposition get it right.

With the simple 3-step linear and non-branching value chains that were analysed so far, other chain variations and aggregations are possible. There are 6 possible orders in which the 3 sectors can be placed, and these yield 15 pairs of distinct chains to aggregate. Generally, when aggregating any of these pairs, a severe distortion in the aggregate decompositions arises, and usually the diagonal values in the decompositions are overestimated. One exception to this general rule is when we have a chain in which $Z_1 = \Delta_2 = \Delta_3$, and the final sector of the two aggregated value chains is identical. But this is a special case that does not arise often in reality.

At an intuitive level, one obvious reason underlying the distortion is the multiple roles played by sectors in the different value chains, i.e., either as an intermediate good provider or as a final good provider. In the examples discussed so far, the value added coefficient in production by Sectors B and C are differentiated according to the sector's position in the value chain. For example, as a final good provider sector C adds $1 - \frac{Z_2}{Z_3}$ of the total value it delivers to final demand, while in value chain 2, sector C adds $1 - \frac{Z_1}{Z_2}$ of the total value it delivers to sector B as intermediate output. The information about this difference is lost by the aggregation of the two chains into one input-output table, and this makes Sector C look like a sector whose output always incorporates a share $1 - \frac{Z_1+Z_2}{Z_2+Z_3}$ value added, regardless of whether the output is a final or an intermediate product. Throughout the Leontief algebra (i.e., **S** = **M**x**L**x**Φ**) that decomposes sectoral value added into the originating sectors of final demand, the homogenisation of the value added coefficient is translated into a major distortion vis-a-vis the "true" decomposition which is not generally obtainable due to the empirical non-observability of the many individual chains.

There is also another factor behind the inflation of the diagonal of matrix **S**. In the matrix multiplication **S** = **M**x**L**x**Φ**, the resulting diagonal will contain elements that are multiplications of the diagonal elements of **Φ** and **L** and **M**. In order to understand the distortion on the diagonal of **S**, one may start by looking at the particular features of the diagonals of matrices **L**, **M** and **Φ** individually. **Φ** is simply the sum of the final demand vectors that characterise individual chains, which is not any source of the distortion.

With regard to the matrix **L** = [I-**A**]$^{-1}$, note that a sufficient condition for the diagonal elements of this matrix to be equal to 1, is that for every non-zero element *i*-*j* (*i* different from *j*) in the underlying matrix [I-**A**], the corresponding element *j-i* is 0. This condition is a technical way of specifying that no production cycles exist. This is indeed the case in both non-aggregated value chains. Thus, in the decomposition of the individual value chains (matrices **S**, **S***, **D** and **D***), the diagonals of the **L** matrix will only contain ones. However, in the aggregate table, production cycles are created that do not exist in the underlying value chains.[2] In the case of the above example, this happens with elements (2, 3) and (3, 2), which both become non-zero in the [I-**A**] matrix for the aggregate table. This leads to a value >1 for elements 2 and 3 on the diagonal of the matrix $[\mathbf{I}\text{-}\mathbf{A}]^{-1}$, as is clearly seen in the above formula for $\mathbf{L}^{\mathbf{agg}}$. This is the second factor that adds to the distortion of the aggregate decomposition.

The second of the aggregation problems, i.e., the influence of production cycles, can be isolated in a theoretical way, by making the value added contribution of each sector constant, irrespective of the stage of the value chain that it appears in. Implementing this for the example used so far, we specify, as was already the case, that Sector A always has $v_A / q_A = 1$, and we specify fixed $v_B / q_B = m_B$ and fixed $v_C / q_C = m_C$ across the two chains. This can be achieved by setting the gross output of sector *i* (= B or C) to $u_{1i} / (1\text{-}m_i)$.

To analyse this case requires setting up the input-output tables from the very beginning. We will not document all steps of this procedure, but instead provide just the following results. First, for the disaggregated value chain A → B → C, we have

D =

1	$1 - m_B$	$(1 - m_C)(1 - m_B)$
0	m_B	$(1 - m_C)m_B$
0	0	m_C

Second, for the aggregated analysis, we have:

$\mathbf{D}^{\mathbf{agg}}$ =

1	$(2m_B^2 - 5m_B + 3)\dfrac{1 - m_C}{3 - m_C - m_B}$	$(2m_C^2 - 5m_C + 3)\dfrac{1 - m_B}{3 - m_C - m_B}$
0	$(2 - m_C)m_B\dfrac{2 - m_B}{3 - m_C - m_B}$	$m_B\dfrac{m_C^2 - 3m_C + 2}{3 - m_C - m_B}$
0	$m_C\dfrac{m_B^2 - 3m_B + 2}{3 - m_C - m_B}$	$(2 - m_B)m_C\dfrac{2 - m_C}{3 - m_C - m_B}$

Under which conditions will the aggregate decomposition provide a proper representation of the value added distribution of the first chain (A → B → C)? In order to answer this question, we solve $(2 - m_B)m_C\frac{2 - m_C}{3 - m_C - m_B} = m_C$, i.e., that the lower-right element of these two matrices is identical. The solution yields only trivial cases: $m_C = 1 \vee m_B = 1 \vee m_C = 0$. These are all triv-

[2] In the Leontief matrix for the aggregated simple chains that we considered so far, at least one cycle will necessarily emerge from any aggregation.

ial because the values are incompatible with the idea of a value chain at all. Thus we conclude that with value-added-coefficients constant for every sector irrespective of the stage of the value chain, the diagonal values of matrix $\mathbf{D}^{\mathbf{agg}}$ will always overestimate the true value added contributions.

2.2. Introducing diagonal cycles

So far, we considered cases where production cycles can only result from aggregation of value chains. This implies that such production cycles are involve more than one sector, i.e., that the diagonal values of the intermediate matrix **U** are 0. In the simple cases that we considered, production cycles comprise two sectors, and they arise as a result corresponding cells below and above the diagonal having values >0. In this section, we briefly look at the case of production cycles that arise within the chain, and which lead to diagonal values of matrix **U** >0. Our aim is not to provide a full analysis of this case, but instead to illustrate that this kind of setup is a game-changer, i.e., the nature of the bias may change as a result of these chain-level cycles.

We analyse the case of aggregating the cycles A → B → C → C and A → C → B → B, and trying to reconstruct the first of these value chains by input-output decompositions. We start by analysing the case where the \$ value added numbers in this cycle are given by Z_1, Δ_2, Δ_3 and Δ_4, respectively. This is similar to the first analytical case above, the only difference being that at the final stage of the chain, $\Delta_3 + \Delta_4$ is added in two stages, instead of just Δ_3 in just one stage. We do not provide details of the calculations, but it can be shown that the condition for whether the diagonal value of **D** of the final sector is overstated does not change. This condition becomes $\Delta_3 + \Delta_4 = 2\left(\frac{\Delta_2^2}{Z_1} + \Delta_2\right)$.

When we specify value chains with fixed value added to output ratios for each sector, the situation does change. Before, we were able to show that for non-trivial values of the value added ratios, the diagonal values of **D** will always be overstated. This is no longer the case. Again, we do not present details of the calculations, which are somewhat tedious, but do produce a relatively clear-cut result, which is displayed in Figure 1.

The figure shows that only the value-added-to-output ratio of the final sector matters in the outcome. Roughly speaking, when this is larger (smaller) than 0.5, the contribution of the final sector to total value added in the chain will be understated (overstated). The vertical line between the two parts of the plane appears to be linear, but in fact it is not. It is a third-degree polynomial that is almost linear in the relevant domains of both axes.

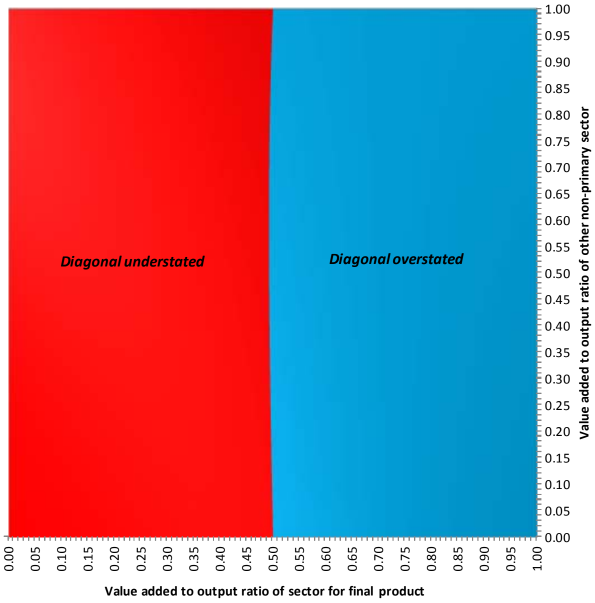

Figure 1. Condition for overstating or understating diagonal values of matrix **D**

2.3. Numerical examples

Table 1a displays a value chain in which the value added that is delivered in each stage declines progressively when final demand is approached. Table 1b displays matrix **S** for this table, i.e., the $ value added decomposition for the chain. By assumption, this is clearly a value chain in which the final stage ("final assembly") adds relatively little value, as case studies have shown for goods such as the iPod. The $6 final demand to sector C is met by the $3, $2, and $1 value added contribution of sector A, B and C respectively.

Table 1a. Input-Output Table for Value Chain 1

	Deliveries to				
	Sector A	Sector B	Sector C	Final demand	Gross Output
Sector A	0	3	0	0	3
Sector B	0	0	5	0	5
Sector C	0	0	0	6	6
Value Added	3	2	1		
Gross Output	3	5	6		

Table 1b. Value Added Decomposition for Value Chain 1

	Sector A	Sector B	Sector C	Row Sum
Sector A	0	0	3	3
Sector B	0	0	2	2
Sector C	0	0	1	1
Column Sum	0	0	6	

Table 2a gives the input-output table for a different value chain, i.e., one where the final good is produced in sector B, whereas sector C provides the intermediate stage. This is a chain A → C → B, whereas the case in Table 1 was a chain A → B → C. Table 2b provides the value added decomposition.

Table 2a. Input-Output Table for Value Chain 2

	Deliveries to				
	Sector A	Sector B	Sector C	Final demand	Gross Out-put
Sector A	0	0	3	0	3
Sector B	0	0	0	6	6
Sector C	0	5	0	0	5
Value Added	3	1	2		
Gross Output	3	6	5		

Table 2b. Value Added Decomposition for Value Chain 2

	Sector A	Sector B	Sector C	Row Sum
Sector A	0	3	0	3
Sector B	0	1	0	1
Sector C	0	2	0	2
Column Sum	0	6	0	

We get to the main point of our argument if we ask what happens if we aggregate these two value chains, into one input-output table. This is done in Table 3a. As we did not change the values of gross output at the respective changes of the chain, the first case analysed in the previous section should apply. We use Table 3a, and all the usual calculations of matrices **D** and **S**, to ask what is the distribution of value added over Value Chain 1. As indicated by Table 3b, the answer would be that, for 1 unit ($6) of final demand for the product that is the final output of Value Chain 1, 0.34 ($2.0625) would be contributed by Sector C (the sector that delivers the good final product of Value Chain 1), 0.16 ($0.9375) would be produced by Sector B, and 0.5 ($3) would be produced by Sector A. Comparing to the actual values in Table 1b, it is clear that these values are not correct, except for Sector A. From the aggregate Table 3b, we would conclude that Sector C has a larger contribution to value added in the chain than Sector B, while in reality (Table 1b), the reverse is the case. The same conclusion arises if we analyse Value Chain 2 by using Table 3. The value added distributions for any chain that we compute using the aggregated table are distorted. The extent of the distortion can be observed clearly by the comparison of Tables 3b and 3c, where the latter is the correct decomposition which is obtained by the aggregation of Tables 1b and 2b, while the former is the incorrect one obtained from the aggregate Table 3a according to the formula **S** = **M**x**L**x**F**.

Table 3a. Input-Output Table for the two value chains aggregated

	Deliveries to				
	Sector A	Sector B	Sector C	Final demand	Gross Output
Sector A	0	3	3	0	6
Sector B	0	0	5	6	11
Sector C	0	5	0	6	11
Value Added	6	3	3		
Gross Output	6	11	11		

Table 3b. Value Added Decomposition Implied by Input-Output Table for the two value chains aggregated

	Sector A	Sector B	Sector C	Row Sum
Sector A	0	3	3	6
Sector B	0	2.0625	0.9375	3
Sector C	0	0.9375	2.0625	3
Column Sum	0	6	6	

Table 3c. Sum of value added decomposition matrices in Tables 1b and 2b

	Sector A	Sector B	Sector C	Row Sum
Sector A	0	3	3	6
Sector B	0	1	2	3
Sector C	0	2	1	3
Column Sum	0	6	6	

Next, we use the m_i values from Value Chain 1 and apply them to Value Chain 2, which will make the gross output levels in that chain equal to 3, 3.6 and 6. Aggregating those two value chains, and then decomposing leads to the results in Table 4. The decompositions are given in Table 4a (based on the aggregate input-output table) and Table 4b (the “correct” aggregation of the two value chains). Despite the fact that the heterogeneity of value added-to-gross output coefficients no longer plays a role, the decomposition based on the aggregate table is still distorted. The presence of cycles in the aggregate table explains this distortion.

Table 4a. Value Added Decomposition Implied by Input-Output Table for the two value chains aggregated, using fixed value added coefficients per sector

	Sector A	Sector B	Sector C	Row Sum
Sector A	0	2.7123	3.2877	6
Sector B	0	2.8923	1.5068	4.4
Sector C	0	0.3945	1.2055	1.6
Column Sum	0	6	6	

Table 4b. Value Added Decomposition for the two value chains, “correct” aggregation, using fixed value added coefficients per sector

	Sector A	Sector B	Sector C	Row Sum
Sector A	0	3	3	6
Sector B	0	2.4	2	4.4
Sector C	0	0.6	1	1.6
Column Sum	0	6	6	

Does the problem go away with larger numbers, i.e., if we aggregate many more value chains? From the simple example value chains that were used so far, this does not seem to be the case. With 3 sectors, 6 different orders are possible, and if we aggregate all these 6 cases altogether, while keeping the value added at each stage of the chain constant, the problem only worsens. In this case (using gross output as in Tables 1 and 2), the value added shares on the diagonal increase to 0.56, while in the underlying value chains that make up the aggregate picture, the share is 1/6 (approximately 0.17).

3. Simulations

The two aims of this Section are, first, to analyse more complicated cases than the simple linear chains of length 3 that we were able to analyse analytically, and, second, to extend the analysis to include a dynamic picture, i.e., to ask whether the static distortions that we identified carry over when we compare changes to value added contributions (or FVAS). We will do this by means of simulation analysis. Before we proceed to lay out the model that we use to simulate the decomposition of value chains, we will take a look at some basic results from the World Input-Output Database (WIOD), in order to set a rough frame for what kind of results to expect from the simulations. Obviously, the WIOD does not provide any information on the underlying disaggregated value chains, so we cannot provide an overview of the actual bias that is implied in the decompositions. But we can collect some basic information on the value added coefficients in WIOD, so that our simulations can be calibrated so that the outcomes resemble at least these "stylised facts".

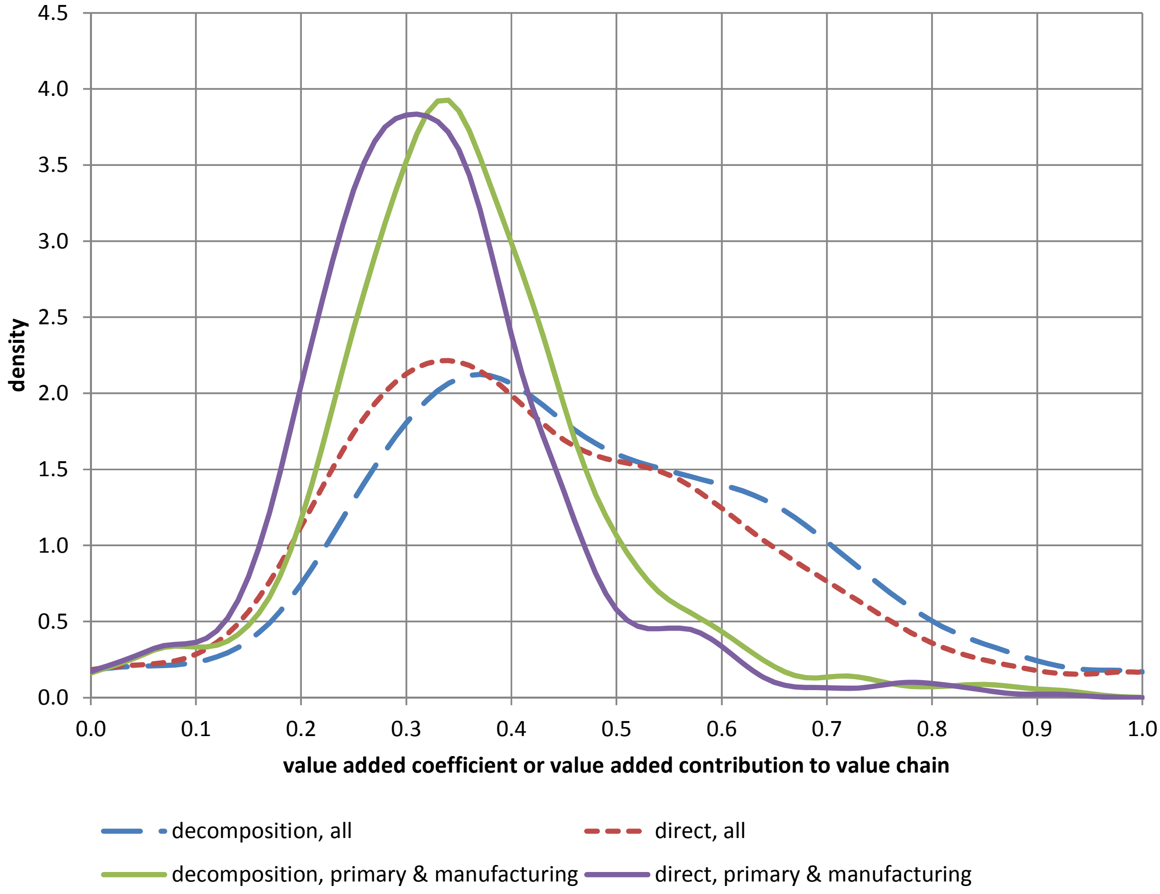

Figure 2. Kernel density plots of matrix elements *d* and *m*, WIOD, 2011

3.1. Empirical impressions from the WIOD

We look at some results of calculations using real-world input-output tables, in particular the WIOD, for 2011. This means we have 40 countries plus a "rest of the world", and 35 sectors for each of those. Hence there are 1435 rows and columns in the matrices **L**, **M** and **D**. First, we look at the values on the diagonal of the matrix **D** = **M**x**L**. Our contention is that under a wide range of circumstances, these tend to over-estimate the true contribution of the sector to the total value added of the value chain of the products produced as final demand by this sector. We cannot test this assertion directly, because we do not have any information on the underlying tables for individual value chains, but we can look at the distribution of these diagonal values of matrix **D**, so that our simulations can be calibrated to reproduce this distribution in approximate form.

In Figure 2, we plot the kernel density distribution of the diagonal values of matrix D (indicated as decomposition), along with the direct value added coefficients (*m*, indicated as direct). The density has been estimated both for all industries (including services) and for the primary and manufacturing sectors separately. The curves for all industries, including services, lie more to the right than the one for primary and manufacturing industries, which indicates the diagonal values that we are interested in tend to be lower in manufacturing and primary industries. The median for the "decomposition" value for the primary and manufacturing industries is 0.45, for all industries it is 0.44.

Figure 3 plots the diagonal values of matrix **M** against those of matrix **D**, for the primary and manufacturing industries. The benchmark in this figure is the 45 degrees line, and it is clearly seen that all values are either on or above this line. The values that are above the 45 degrees line are the result of production cycles in the input-output table, which is one of the factors behind the over-statement of the diagonal values of the **D** matrix. We clearly see that these production cycles are important for many sectors, as many observations are well above the 45 degrees line.

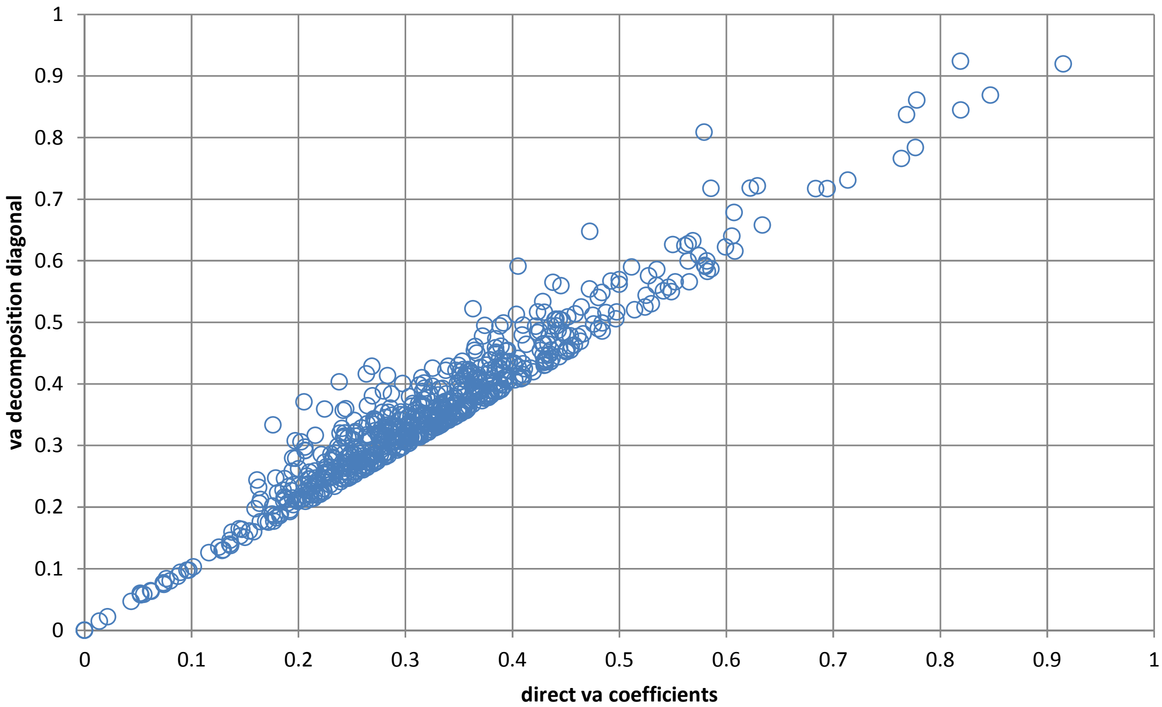

Figure 3. Matrix elements *m* vs *d* (diagonals), WIOD, 2011

3.2. Simulations

We implement simulations in Visual Basic in Excel 2010. In each simulation, 150 value chains are built up in a random way, by first picking a primary sector (defined as a sector that uses no intermediate inputs) and then picking a number of non-primary sectors to form a chain. We limit the experiments to cases where all value chains are a sequence without branches (i.e., so-called "snakes"). The length of the chains is normally distributed, and the mean and standard deviation of the normal distribution is taken as an input-parameter in the simulations. We use an average chain length of 4 with standard deviation 2.

The simulated chain can directly involve production cycles, i.e., the same sector may arise more than once in a single chain, which will lead to the diagonal of the input-output matrix being populated. Also the aggregation of chains can lead to production cycles, and these are "off-diagonal". There are 5 primary sectors, and these sectors only occur as primary, i.e., at the start of the chain. A total of 30 sectors is used as non-primary, which implies that the input-output tables that we will build are 35x35 sectors.

The probability of each sector to be picked for the next stage of the chain is, in principle, uniformly distributed. However, we also partition the sectors into groups, mimicking "countries". In this case, we set a probability P, which we call the domestic bias. The uniform probability is then multiplied by $1 - P$ for sectors that belong to a "foreign country". The primary use of the domestic bias parameter is to run dynamic experiments, in which we compare a situation with high domestic bias with one with low bias. For the dynamic experiments, we also vary chain length. When we use a domestic bias in the simulations, the non-primary sectors are divided into 5 countries, each of which includes exactly one primary sector.

There is a choice between two regimes that determine the amount of value added generated at each stage of the chain. In Regime 1, value added contributions are defined as $ values, and can consists of three components: one specific to the sector (the sector always contributes this amount), one specific to the stage (the fixed $ value is always contributed by the stage), and a chain specific contribution (a fixed $ value fixed throughout the chain). The value added produced at every stage of the chain is the sum of these 3 components, and each of the components is drawn from a uniform distribution for which we define the boundaries. In Regime 2, the value-added-to-gross-output coefficient is drawn from a normal distribution, with the mean dependent on the sector, and a fixed standard deviation. The sector-dependent mean is fixed over the simulation, and itself drawn from a uniform distribution. Hence in this regime, the value-added-to-output ratio shows some heterogeneity over the chains, but has a strong sector-specific element.

For each value chain generated in the simulation, we standardise all value added values to generate a final output value that is equal to a random value drawn from the uniform distribution [7,13]. The value added and gross output numbers at each of the stages are then entered in a matrix **S****, which will contain the true value added decomposition, sector-by-sector, and into an input-output system (**U**, **F** and **Q**) that can be used to perform the decompositions used in Section 2. We then calculate **S – S****, where **S** is based on the input-output system, as a way to assess how well the decompositions capture the true value added structure.

3.2.1. Static simulations

Table 5 documents some descriptive results of 5 basic simulations. The first 3 columns present results for value added regime 1, the 4th and 5th column for value added regime 2. The table provides statistics on the distributions of the variables covered in Figures 2 and 3. These are intended to assess how well the simulation results reproduce the stylised facts of the real-world input-output tables, even though the real-world tables contain many more rows/columns than our simulated tables. In the simulations used for Table 5, there is no domestic bias in the construction of value chains ($P = 0$). All statistics in the table are averages over 5 runs with different random seeds, and the numbers between brackets are standard deviations over these 5 runs.

The *m* statistic in the table refers to the value-added-to-output ratios of each sector, i.e., what is called “direct” in Figure 2. The statistic referred to as *diag* is conceptually equal to the diagonal values of the **D** matrix, i.e., what is called “decomposition” in Figure 2. Note, however, that we cannot calculate the **D** matrix for our “true” values in the simulation (**D** = **M**x**L**, and there is no such thing as an **L** matrix for our true values). Hence we calculate a “true **D** matrix” by dividing each cell of the matrix **S*** by its column sum, and this is what is called *diag* in the table. Finally, we have a statistic called *diff*, which is *diag – m*, and which refers to Figure 3, where it is the vertical distance between a dot and the 45 degrees line.

In the true distribution (Figure 2), the median values for *m*, *diag* and *diff* are slightly smaller than the average, i.e., the distribution is slightly skewed to the right. This is reproduced in the simulations with value added regime 2, but not regime 1, where the median is slightly larger than the mean. For columns 1, 4 and 5, the values for the median, average and standard deviation of *diag* and *m* are roughly of the same order of magnitude as for the empirical distributions in Figure 2, if we focus on the manufacturing and primary sectors alone. For columns 2 and 3, this is the case for all indicators, except the standard deviations, which are somewhat smaller than observed empirically.

The 3 simulations with value added regime 1 differ with respect to the source of variation that they introduce. Column 1 only introduces sectoral variation in the value added contributions. Column 2 has sectoral- and stage-specific variation, while column 3 has variation of all 3 sorts, i.e., also introduces chain-specific variation. The effect of introducing non-sector-specific variation in the value added contributions (i.e., either stage- or chain-specific) is that the standard deviation of the various indicators declines (rows Std). This is as expected, because the sectors may occur in different stages of the chain, and therefore, a stage-specific variation will tend to draw the sectors towards an average value. The impact of the chain-specific variation is limited, because the total final demand gets normalised into a limited (random range). Nevertheless, adding this component still reduces the standard deviations. A similar effect, although to a lesser extent, is observed in the 2 columns for value added regime 2. The first of these columns has standard deviation = 0, i.e., the sector-specific value added to output ratio is fixed throughout the simulation. In the last column, the standard deviation of the sectoral value added to output ratio is 0.3 (mean is fixed at 0.4).

Table 5. Basic simulation results (no domestic bias)

	Regime 1 Only Sector		Regime 1 Sector + Stage		Regime 1 All components		Regime 2 std=0		Regime 2 std = 0.3	
Median *m*	0.344	(0.024)	0.318	(0.008)	0.304	(0.009)	0.388	(0.043)	0.397	(0.040)
Average *m*	0.313	(0.018)	0.310	(0.007)	0.312	(0.008)	0.401	(0.023)	0.427	(0.020)
Std *m*	0.133	(0.010)	0.089	(0.013)	0.074	(0.014)	0.188	(0.011)	0.144	(0.017)
Median *d*	0.362	(0.021)	0.341	(0.018)	0.319	(0.014)	0.397	(0.045)	0.415	(0.043)
Average *d*	0.326	(0.019)	0.324	(0.010)	0.324	(0.009)	0.409	(0.023)	0.439	(0.024)
Std *d*	0.139	(0.010)	0.092	(0.013)	0.078	(0.015)	0.190	(0.012)	0.145	(0.017)
Median *diff*	0.008	(0.001)	0.009	(0.001)	0.008	(0.001)	0.005	(0.001)	0.007	(0.001)
Average *diff*	0.013	(0.002)	0.014	(0.003)	0.013	(0.001)	0.008	(0.002)	0.012	(0.004)
Std *diff*	0.013	(0.003)	0.012	(0.004)	0.011	(0.003)	0.008	(0.002)	0.013	(0.006)

Notes: In Regime 1, sector contribution drawn from [0,40], Stage and Chain both drawn from [0,20], in Regime 2, Mean = 0.4.

Table 6 presents statistics on the distortions of the value added distribution, i.e., on the difference **S – S****. In this table, we present results for five different values of the internationalisation parameter *P*. On the left, we have a completely internationalised case, where country borders play no role ($P = 0$), and this goes, by intermediate values of *P*, to the opposite case, where all value chains are concentrated in a single country ($P = 1$). Again, the table reports averages over five different random seeds, with standard deviations over seeds reported in brackets.

The first part of the table, the "overall distortion" part, sums all cells of the matrix **S – S**** that have a positive deviation, and divides by the total value added in all chains in the simulations. Note that the total value added is the same in matrix **S** and matrix **S****, which implies that the sum of deviations in matrix **S – S**** is 0. Hence the "overall distortion" statistic is matched by a negative counterpart. In other words, the sum of absolute deviations in the matrix **S – S**** is twice the magnitude of the "overall distortion" statistics.

Table 6. Simulation results: various measures of distortion of the value added distribution

	Country compartmentalisation					
Overall distortion	**0.00**	**0.25**	**0.50**	**0.75**	**0.95**	**1.00**
Regime 1, sector	0.340	0.326	0.325	0.312	0.238	0.108
	(0.014)	(0.013)	(0.016)	(0.021)	(0.026)	(0.006)
Regime 1, sector + stage	0.338	0.330	0.335	0.314	0.252	0.117
	(0.016)	(0.006)	(0.009)	(0.018)	(0.018)	(0.009)
Regime 1, all	0.348	0.335	0.336	0.318	0.248	0.115
	(0.015)	(0.017)	(0.006)	(0.008)	(0.021)	(0.009)
Regime 2, std=0	0.251	0.242	0.250	0.230	0.173	0.089
	(0.018)	(0.013)	(0.012)	(0.017)	(0.026)	(0.002)
Regime 2, std=0.3	0.280	0.273	0.277	0.258	0.202	0.123
	(0.020)	(0.017)	(0.016)	(0.018)	(0.022)	(0.008)
Diagonal positive distortion						
Regime 1, sector	0.152	0.138	0.120	0.114	0.079	0.076
	(0.023)	(0.019)	(0.014)	(0.019)	(0.022)	(0.014)
Regime 1, sector + stage	0.154	0.145	0.130	0.114	0.096	0.080
	(0.026)	(0.021)	(0.016)	(0.028)	(0.015)	(0.020)
Regime 1, all	0.154	0.133	0.137	0.108	0.097	0.064
	(0.035)	(0.025)	(0.017)	(0.019)	(0.013)	(0.012)
Regime 2, std=0	0.010	0.010	0.013	0.010	0.012	0.014
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
Regime 2, std=0.3	0.047	0.043	0.045	0.040	0.033	0.030
	(0.010)	(0.006)	(0.013)	(0.003)	(0.006)	(0.006)
Diagonal negative distortion						
Regime 1, sector	-0.011	-0.009	-0.009	-0.012	-0.025	-0.027
	(0.007)	(0.005)	(0.010)	(0.005)	(0.009)	(0.010)
Regime 1, sector + stage	-0.013	-0.013	-0.014	-0.023	-0.030	-0.034
	(0.003)	(0.004)	(0.007)	(0.009)	(0.010)	(0.009)
Regime 1, all	-0.012	-0.014	-0.012	-0.019	-0.035	-0.030
	(0.005)	(0.006)	(0.006)	(0.005)	(0.008)	(0.005)
Regime 2, std=0	-0.013	-0.017	-0.016	-0.022	-0.029	-0.034
	(0.002)	(0.004)	(0.005)	(0.006)	(0.007)	(0.007)
Regime 2, std=0.3	-0.051	-0.049	-0.056	-0.061	-0.054	-0.058
	(0.008)	(0.006)	(0.011)	(0.004)	(0.017)	(0.007)
FVAS overall distortion						
Regime 1, sector	0.148	0.131	0.149	0.215	0.500	--
	(0.020)	(0.012)	(0.039)	(0.042)	(0.059)	--
Regime 1, sector + stage	0.126	0.129	0.153	0.195	0.408	--
	(0.009)	(0.021)	(0.022)	(0.038)	(0.034)	--
Regime 1, all	0.124	0.132	0.145	0.191	0.420	--
	(0.016)	(0.020)	(0.017)	(0.015)	(0.049)	--
Regime 2, std=0	0.098	0.102	0.130	0.192	0.390	--
	(0.010)	(0.019)	(0.015)	(0.027)	(0.038)	--
Regime 2, std=0.3	0.174	0.169	0.198	0.239	0.380	--
	(0.017)	(0.018)	(0.028)	(0.030)	(0.048)	--

The overall distortion declines with the value of the parameter *P*. It declines slightly for values of P up to 0.75, and then declines sharply. Thus, the overall distortion is smallest for the case where value chains are completely domestic. Between the parameter sets, the overall distortion is larger in the simulations with regime 1 than those with regime 2, although they are still sizeable in regime 2. In regime 2, the overall error increases when the standard deviation goes up from 0 to 0.9. In regime 1, the overall error does not vary much between the 3 simulations.

The next two blocks of the table look at the diagonal values of **S – S****, and here we consider positive and negative deviations separately. The sum of the positive or negative cells are divided by total value added on the diagonal in the true decomposition (matrix **S****). The analytical cases above suggest that the values on the diagonal are biased upward, especially so in value added regime 1. This is clearly confirmed in all simulations with value added regime 1. In this case, the positive diagonal error is around 15% for the case when country borders do not matter, and this declines to about half (7%) when value chains are concentrated domestically. The negative diagonal distortion, on the other hand, is very much smaller in this regime, and it does not vary much with the degree of internationalisation. With this regime, the size of the distortion on the diagonal grows slightly with adding more components of variance of the value added contributions.

In line with the conclusions from the analytical exposition, the distortion on the diagonal is much smaller for simulations with value added regime 2. This regime clearly emerges as the one that is most "favourable" to input-output analysis. In addition, the small distortion that is observed in regime 2 is fairly evenly distributed in the negative and positive domain (with the negative domain slightly dominating), and it does not vary much with the level of internationalisation. However, when the standard deviation is set to 0.3 (last row of each block), the distortion on the diagonal grows, even if it never comes close to the values observed with regime 1.

Finally, we look at the FVAS statistic. Since FVAS includes both the diagonal value and some of the off-diagonal values of the matrix **S**, it will combine features of the diagonal and off-diagonal distortions. The statistic given in Table 6 is the average of the sum of absolute differences in FVAS for each of the 30 non-primary sectors, between **S** and **S****, divided by the true average value of FVAS over the 30 sectors.

There are two common trends in the FVAS distortion results. First, the (relative) distortion increases when the economy becomes less internationalised (i.e., the numbers increase from left to right in the table). Second, the level distortion differs between regimes 1 and 2, but also varies greatly within regime 2, depending on the standard deviation. The distortions are sizeable, varying between about 10% and 17% for the completely globalised case, and up to 50% for cases close to completely domestic value chains (obviously, FVAS = 0 when *P* = 1, so our statistic is not defined for the last column in the table).

3.2.2. Dynamic simulations

We now turn to investigating the dynamic nature of the distortion and biases that were observed so far, and focus on FVAS. This addresses the conclusions reached in part of the input-output literature (Timmer et al., 2014; Los et al., 2014) about an increase in so-called international fragmentation (FVAS). For example, Los et al. (2014) observe that over the 1995 – 2008, FVAS goes up significantly, and for the large majority of (manufacturing) sectors in the WIOD database. This is seen as evidence of an increasing importance of international or global value chains.

Our results so far suggest that the FVAS statistics that Los et al. (2014) use for any of the individual two years that they use, are probably biased. Specifically for FVAS, this would be the case irrespective of whether the real world operated as in our value added regime 1 or regime 2. However, if such a bias exists in both years that Los et al. (2014) observe, could it be that the change between those 2 years is still un-biased? In other words, is the bias static, or dynamic? This is the issue that we address in this section.

In the dynamic setting, a new aspect of the input-output decompositions comes up. This is the fact that the inverse Leontief calculation has the effect of evening out many of the distributions that are found in the underlying true value chains. This is most obvious in the fact that the true matrix **S**** has relatively many cells that are equal to zero, while the matrix **S**, based on the inverse Leontief calculations typically has no zero-cells at all. In other words, if we look at a single column of the matrix **S** (the distorted value added matrix), the values in this column suggest that all other sectors in the system contribute to the value added in the demand delivered by the column-sector. In the true matrix **S****, however, there is a much more limited number of sectors that contributes to this value added, as there are many zeros in the column. This phenomenon, which (of course) also played a role in the static simulations, has an important consequence in the dynamic context, which we will discuss below.

The setup of our simulation experiment is as follows. For a given random seed, we carry out two simulations: one with a domestic bias equal to 0.75, and one with domestic bias equal to 0.25. Because the random seed is fixed between the two runs, the length of the chains, and the components that build up the value added in the chain, does not vary between the two runs. This implies that the value added contributions of the sectors (whether they are specified in $ values, or in value-added-to-output ratios) do not vary between the runs, which makes the two experiments comparable in this respect. The only thing that varies, is the tendency for chains to become internationalised.

Then we compare FVAS between the two runs, both for the true value added matrix **S****, and for the distorted matrix **S**, and ask whether the true change in FVAS is well approximated by the change in FVAS that we see in the matrix based on the input-output decompositions (**S**). The change from domestic bias = 0.75 to domestic bias = 0.25 is intended to mimic a globalisation shock, similar to what is found by Los et al. (2014). We expect that as a result of this shock, FVAS will go up.

The basic results are displayed in Figures 4a and 4b, which plot the true change in FVAS against the decomposed change in FVAS. Both figures collect data over 5 different random seeds per parameter setting. The observations are the 30 non-primary sectors, so that there are 150 observations in each figure. The black (non-steepest) line is the regression line fitted through the observations, the other line is the 45 degrees line. If changes in FVAS in the decomposed matrix **S** would be a reliable approximation of true changes in FVAS, we would expect the regression line to be close to the 45 degrees line, and the points to be spread in a close band around this line.

Figure 4a gives the results for value added regime 1, with only sectoral variation in value added contributions. Figure 4b gives results for value added regime 2, standard deviation = 0. For both cases, we observe a clear bias in the decomposed FVAS-changes. The other regimes of Tables 5 and 6 are not documented here in order to save space, but they yield similar results.

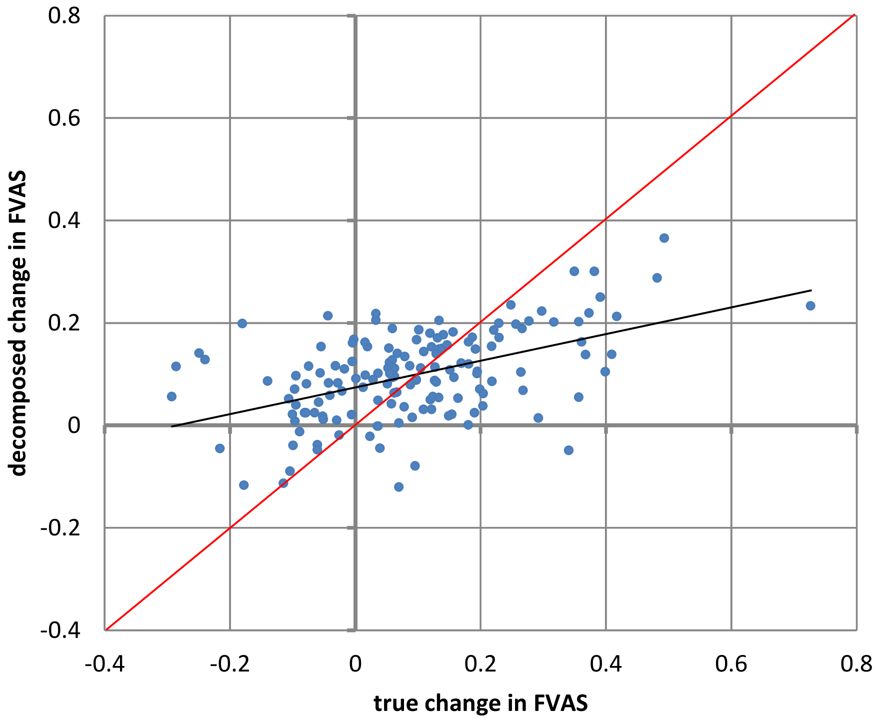

Figure 4a. True FVAS vs. decomposed FVAS, value added regime 1, only sector variation

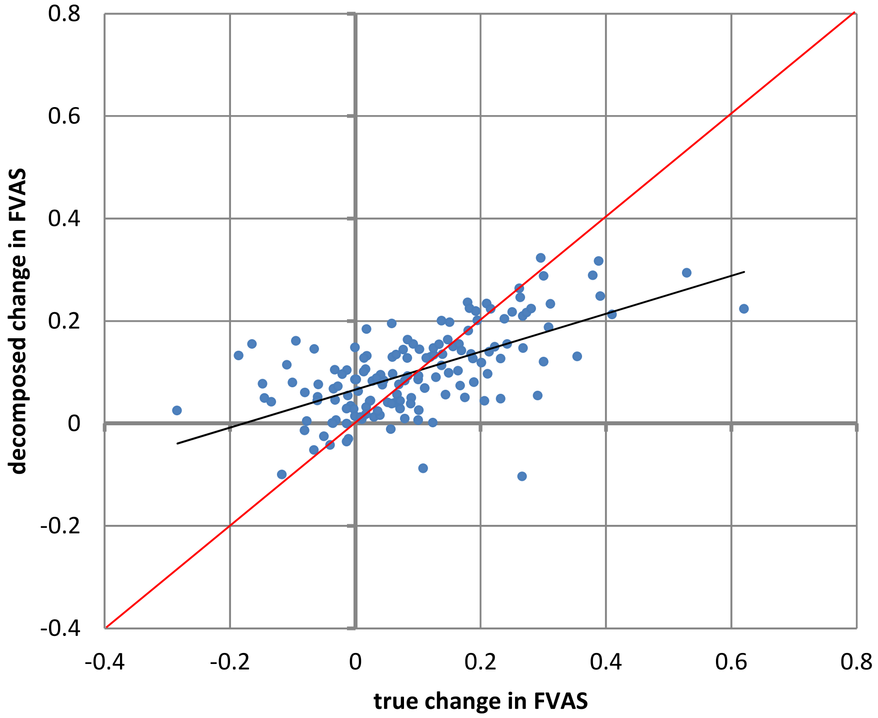

Figure 4b. True FVAS vs. decomposed FVAS, value added regime 2, standard deviation=0

The most obvious feature in both figures is that the regression lines have slopes that are clearly smaller than 1 (in Figure 4a, the slope is 0.26, in Figure 4b, it is 0.37). This means that large (small) true changes in FVAS are under- (over-) estimated by the input-output decompositions, as is also clearly seen by the position of the observations relative to the 45 degrees line. In addition to this bias, the points are also spread out fairly widely around the regression line, i.e., there is a large amount of noise in the approximation of the true changes in FVAS. The R^2 of the regression line in Figure 4a is 0.25, in Figure 4b it is 0.38.

The bias in the approximated FVAS changes is the result of a process that can be seen as a regression to the mean. We already pointed to the tendency of the decomposed value added matrix **S** to even out the value added contributions within a column (and across rows). This is essentially a characteristic of the static value added decomposition, and it holds both for domestic bias setting 0.25 and 0.75. As a result of this static tendency to even out differences in FVAS between the sector, the dynamic changes of FVAS are much smaller in the decomposed matrices **S** than in the true matrices **S****.[3]

It is as if the decomposition matrix **S** "smears out" a spectrum of widely differing colours (found in the true matrix **S****) into various shades of grey. Although the results with domestic bias = 0.25 yield a darker shade of grey than the results with domestic bias = 0.25, the difference between these two shades of grey is much smaller than the differences in colour that are found when the true matrix **S**** is used. Switching from the colour analogy back to the actual results, this tendency is clearly observed in Figures 4a&b by looking at the range of values found on the horizontal axis and comparing it with the range on the vertical axis. The range on the vertical axis is much smaller than the range on the horizontal value. As a result of the fact that the wide range on the horizontal axis is squeezed into a narrow range on the vertical axis, the regression line gets slope < 1.

It is easy to see what this implies for the estimates of changes in FVAS that Los and Timmer present. They find an increase in FVAS that is found pretty much across the board, i.e., both for sectors that initially had low FVAS, and for sectors that initially had high FVAS. Our results suggest that the changes of the initially low values may be over-estimated, and hence that they may not have increased their fragmentation (FVAS) at all. On the other hand, the initially FVAS values have probably grown even more than the results by Timmer and Los suggest. Hence the general increase in international fragmentation that they find may well be less general than they suggest.

The regression to the mean tendency occurs somewhat independently of the (static) biases in the matrix **S** that we analysed extensively above. It occurs, in a relatively strong way, even in cases where the static biases of the diagonal values of matrix **S** are minimal (Figure 4b).

4. Conclusions and discussion

Our analysis shows, both analytically and by means of simulation analysis, that the input-output decomposition methods that are applied to provide evidence about international fragmentation in value chains, are grossly distorting the true underlying value added distribution in an aggregated set of value chains. We pointed to two different forms and sources of such biases.

First, focusing on the final sector, i.e., the sector that delivers the final product of the chain, we find that there is a tendency for the contribution of this sector to be over-estimated. This implies, among other things, that the results of case studies (e.g., on the iPod or mobile phones) that show a small contribution of the final (manufacturing) sector that delivers the product, cannot be com-

[3] In the extreme case of comparing a case with domestic bias = 1 to a case with domestic bias = 0 (or any positive number), the static regression to the mean effect becomes equal to the dynamic effect. In the case of domestic bias, FVAS = 0 for all sectors. Hence FVAS for the domestic bias = 1 (or >0) case is equal to the change in FVAS.

pared to aggregate studies based on input-out tables, because the input-output approach may well over-estimate the distribution of the final sector.

This bias is strongest where the $ value added contribution is specific to the sector that produced it (rather than the stage of the chain it appears in), and it is virtually absent if sectors are characterised by a fixed value-added-to-output ratio, irrespective of which chain they contribute to, and irrespective of the stage of the chain the sector appears in. Hence, such a "scenario" emerges as the one that is least problematic for input-output analysis of value chains.[4] But even in this scenario, the distortion of the value added contribution of non-final sectors is very large as well. In addition, when the value-added-to-output ratio within a sector shows variation over the various chains that the sector participates in, the distortion grows considerably. Again, we observe that for high (low) value-added-to-output ratios, the distortion is negative (positive).

We are unaware of any systematic evidence on whether the fixed $ value scenario or the fixed value-added-to-gross-output ratio scenario is empirically more relevant. Before the input-output decompositions of fragmentation of value chains are widely adopted as an acceptable methodology, we call for an effort to collect such evidence.

Second, our simulation analysis showed that also dynamic comparisons of specific results of the static decompositions tend to be biased. We focused on the foreign value added share (FVAS) statistic that is used in Los et al. (2014) and Timmer et al. (2014). They argue that value chains tend to become more internationally fragmented since 1995 (FVAS tends to increase). Our results point to a bias in the changes in FVAS when calculated using the Los & Timmer method. The bias is a result of a process that we describe as regression to the mean. The underlying static decompositions, and in particular the inverse Leontief matrix used in these decompositions, tends to even out the value added distribution in a rather strong way, leading to much less variation of FVAS values in the decompositions as compared to the true values. In the dynamic context, this effect is reinforced, leading to the bias that we observe: small (large) changes in FVAS tend to be over- (under-) estimated.

Although we focused on the issue of global value chains and international fragmentation, it seems likely that the sources of the biases and distortion that we pointed to (the use of fixed value-added-to-output ratios, the emergence of cycles as a result of aggregation, and the regression to the mean effect in the inverse Leontief matrix) have consequences for other applications of input-output analysis as well. For example, the calculation of so-called footprints (e.g., Al-

[4] It may be noted that the case that is most "favourable" for input-output analysis, is the "scenario" where the value-added-to-output ratio is constant over all chains in which the sector participates, but differs between sectors. This is when the distortion of the value added contribution of the final sector is minimal. But in such a case, the micro level evidence should indeed be a good approximation of the aggregate picture as well. Because the value-added-to-output ratio does not differ (much) between chains within a sector, it is enough to analyse only a few cases. However, the fact that input-output analysis and the micro level case studies tend to disagree on the value added contribution of the final sector, makes it unlikely that this is a relevant scenario.

samawi et al., 2014), which, broadly speaking, measures how consumption instead of production leads to phenomena such as CO_2 emissions, is also likely to suffer from the problems that we pointed to. The same holds for methods that attempt to calculate trade in value added (e.g., Johnson and Nogueira, 2012).

References

Ali-Yrkkö, J., Rouvinen, P., Seppälä, T. and P. Ylä-Anttila, (2011), Who Captures Value in Global Supply Chains? Case Nokia N95 Smartphone, *Journal of Industrial Competition and Trade*, vol. 11, pp. 263–278.

Alsamawi, A., Murray, J. & M. Lenzen, (2014), The Employment Footprints of Nations. Uncovering Master-Servant Relationships, *Journl of Industrial Ecology*, vol. 18, pp. 59-70.

Dedrick, J.,Kraemer, K.L. and G. Linden, (2009), Who profits from innovation in global value chains?: a study of the iPod and notebook PCs, *Industrial and Corporate Change*, vol. 19, pp. pp. 81–116.

Johnson, R. C. and G. Noguera (2012) “Accounting for Intermediates: Production Sharing and Trade in Value Added”. *Journal of International Economics*, 86(2), 224-236.

Los, B., Timmer, M. and G. de Vries, (2014), How Global Are Global Value Chains? A New Approach to Measure International Fragmentation, *Journal of Regional Science*, 2014, preprint.

Miller, R.E. and P. D. Blair, (2009), Input-Ouput Analysis. Foundations and Extensions. 2nd Edition, Cambridge UK: Cambidge University Press.

Shin, N., Kraemer, K.L. and J. Dedrick, (2010), Value Capture in the Global Electronics Industry: Empirical Evidence for the ‘Smiling Curve’ Concept, paper presented at the Industry Studies Conference, May 6-7, 2010.

Timmer, M., Erumban, A.A., Los, B., Stehrer, R. And G. de Vries, (2014), Slicing up Global Value Chains, *Journal of Economic Perspectives*, vol. 28, pp. 99–118.

The UNU-MERIT WORKING Paper Series

2014-01 *The medium-term effect of R&D on firm growth* by Marco Capasso, Tania Treibich and Bart Verspagen
2014-02 *Diverse and uneven pathways towards transition to low carbon development: The case of diffusion of solar PV technology in China*Michiko Iizuka
2014-03 *User innovators and their influence on innovation activities of firms in Finland* by JariKuusisto, MerviNiemi and Fred Gault
2014-04 *Migration, remittances and household welfare in Ethiopia* by Lisa Andersson
2014-05 *Path-breaking directions of nanotechnology-based chemotherapy and molecular cancer therapy* by Mario Coccia and Lili Wang
2014-06 *Poor trends - The pace of poverty reduction after the Millennium Development Agenda*Richard Bluhm, Denis de Crombrugghe, Adam Szirmai
2014-07 *Firms' adoption of international standards: Evidence from the Ethiopian floriculture sector* by MuluGebreeyesu
2014-08 *School choice, segregation, and forced school closure*by Cheng Boon Ong andKristof De Witte
2014-09 *Gender difference in support for democracy in Sub-Saharan Africa: Do social institutions matter?*by MatyKonte
2014-10 *Why are women less democratic than men? Evidence from Sub-Saharan African countries* by Cecilia García-Peñalosa and MatyKonte
2014-11 *Tipping points? Ethnic composition change in Dutch big city neighbourhoods* by Cheng Boon Ong
2014-12 *Technology life cycle and specialization patterns of latecomer countries. The case of the semiconductor industry* by Giorgio Triulzi
2014-13 *Patents as quality signals? The implications for financing constraints on R&D* by Dirk Czarnitzki, Bronwyn H. Hall and Hanna Hottenrott
2014-14 *Assessment of effectiveness of Chinese aid in competence building and financing development in Sudan* by SamiaSatti Osman Mohamed Nour
2014-15 *Education, training and skill development policies in Arab Gulf countries: Macro-micro overview* by SamiaSatti Osman Mohamed Nour
2014-16 *Structure of labour market and unemployment in Sudan* by SamiaSatti Osman Mohamed Nour
2014-17 *Overview of knowledge transfer in MENA countries - The case of Egypt* by SamiaSatti Osman Mohamed Nour
2014-18 *The impact of ICT in public and private universities in Sudan*by SamiaSatti Osman Mohamed Nour
2014-19 *End-user collaboration for process innovation in services: The role of internal resources* by Mona Ashok, Rajneesh Narula and Andrea Martinez-Noya
2014-20 *Public investment and regional politics: The case of Turkey* by Mehmet Guney Celbis, Denis de Crombrugghe and Joan Muysken
2014-21 *Infrastructure and the international export performance of Turkish regions* by Mehmet Guney Celbis, Peter Nijkamp and Jacques Poot
2014-22 *Discovering and explaining work-family strategies of parents in Luxembourg* by Nevena Zhelyazkova
2014-23 *Parental leave take up and return to work of mothers in Luxembourg: An application of the model of nested dichotomies* by Nevena Zhelyazkova

2014-24 *Millennium Development Goals: Tool or token of global social governance?* by Mueid Al Raee, Elvis Amoateng, Elvis Korku Avenyo, Youssef Beshay, Mira Bierbaum, Charlotte Keijser and Rashmi Sinha

2014-25 *One Europe or several? Causes and consequences of the European stagnation* by Jan Fagerberg and Bart Verspagen

2014-26 *The harmony of programs package: Quasi-experimental evidence on deworming and canteen interventions in rural Senegal* by Théophile Azomahou, Fatoumata Diallo and Wladimir Raymond

2014-27 *Country Terms of Trade 1960-2012: Trends, unit roots, over-differencing, endogeneity, time dummies, and heterogeneity* by Thomas Ziesemer

2014-28 *The structure and comparative advantages of China's scientific research - Quantitative and qualitative perspectives* by Lili Wang

2014-29 *Transition to knowledge-based economy in Saudi Arabia* by Samia Satti Osman Mohamed Nour

2014-30 *Challenges and opportunities for transition to knowledge-based economy in Arab Gulf countries* by Samia Satti Osman Mohamed Nour

2014-31 *Migration of international students and mobilizing skills in the MENA Region* by Samia Satti Osman Mohamed Nour

2014-32 *Beyond product innovation; improving innovation policy support for SMEs in traditional industries* by René Wintjes, David Douglas, Jon Fairburn, Hugo Hollanders and Geoffrey Pugh

2014-33 *The impact of innovation support programmes on SME innovation in traditional manufacturing industries: an evaluation for seven EU regions* by Dragana Radicic, Geoffrey Pugh, Hugo Hollanders and René Wintjes

2014-34 *Beliefs dynamics in communication networks* by Théophile T. Azomahou and Daniel C. Opolot

2014-35 *Stability and strategic diffusion in networks* by Théophile T. Azomahou and Daniel C. Opolot

2014-36 *Epsilon-stability and the speed of learning in network games* by Théophile T. Azomahou and Daniel C. Opolot

2014-37 *Afghan unaccompanied minors in the Netherlands: Far away from home and protected?* by Carla Buil and Melissa Siegel

2014-38 *Multinational production and trade in an endogenous growth model with heterogeneous firms* by Hibret B. Maemir and Thomas Ziesemer

2014-39 *The political economy of research and innovation in organic photovoltaics (OPV) in different world regions* by Serdar Türkeli and René Kemp

2014-40 *Towards the societal system of innovation: The case of metropolitan areas in Europe* by Serdar Türkeli and René Wintjes

2014-41 *To return permanently or to return temporarily? Explaining migrants'* intentions by Özge Bilgili and Melissa Siegel

2014-42 *Catching up and lagging behind in a balance-of-payments-constrained dual economy* by Alejandro Lavopa

2014-43 *An introduction to the economics of rare earths* by Eva Bartekova

2014-44 *The unequal effect of India's industrial liberalization on firms' decision to innovate: Do business conditions matter?* By Maria Bas and Caroline Paunov

2014-45 *Insurgents in motion: Counterinsurgency and insurgency relocation in* Iraq by Pui-hang Wong

2014-46 *Successive leadership changes in the regional jet industry* by Daniel Vertesy
2014-47 *Demand, credit and macroeconomic dynamics: A microsimulation model* by Huub Meijers, Önder Nomaler and Bart Verspagen
2014-48 *Accessibility analysis as an urban planning tool: Gas station location* by D.A. Escobar , C. Cadena-Gaitán, F.J. Garcia
2014-49 *China's economic embrace of Africa - An international comparative perspective* by Tobias Broich and Adam Szirmai
2014-50 *Mapping regional social enterprise ecosystems in India: Framework and indicators* by Lina Sonne
2014-51 *Does shelter assistance reduce poverty in Afghanistan?* By Craig Loschmann, Christopher R. Parsons and Melissa Siegel
2014-52 *How important is innovation? A Bayesian factor-augmented productivity model on panel data* by Georges Bresson, Jean-Michel Etienne and Pierre Mohnen
2014-53 *Does too much work hamper innovation? Evidence for diminishing returns of work hours for patent grants* by Mehmet Güney Celbi¸ Serdar Türkeli
2014-54 *Globalization, the rise of biotechnology and catching up in agricultural innovation: The case of Bt technology in India* Michiko Iizuka and Ajay Thutupalli
2014-55 *Where are innovation indicators, and their applications, going?* by Fred Gault
2014-56 *Productivity in services in Latin America and the Caribbean* by Elena Arias-Ortiz, Gustavo Crespi, Alejandro Rasteletti and Fernando Vargas
2014-57 *Optimal public investment, growth, and consumption: Fresh evidence from African countries* by Augustin Kwasi Fosu, Yoseph Yilma Getachew and Thomas H.W. Ziesemer
2014-58 *International R&D alliances by firms: Origins and development* by Rajneesh Narula and Andrea Martínez-Noya
2014-59 *Appropriability mechanisms, innovation and productivity: Evidence from the UK* by Bronwyn H. Hall and Vania Sena
2014-60 *The size of patent categories: USPTO 1976-2006* by François Lafond
2014-61 *Asset recombination in international partnerships as a source of improved innovation capabilities in China* by Simon C. Collinson and Rajneesh Narula
2014-62 *The viability of sustained growth by India's MNEs: India's dual economy and constraints from location assets* by Rajneesh Narula
2014-63 *The effect of unconditional cash transfers on adult labour supply: A unitary discrete choice model for the case of Ecuador* by Andrés Mideros and Cathal O'Donoghue
2014-64 *Giving in South Africa: Determining the influence of altruism, inequality aversion and social capital* by Nyasha Tirivayi
2014-65 *The impact of food transfers for people living with HIV/AIDS: Evidence from Zambia* by Nyasha Tirivayi and Wim Groot
2014-66 *Size and economies of scale in higher education and the implications for mergers* by Nyasha Tirivayi, Henriette Maasen van den Brink and Wim Groot
2014-67 *Widowhood and barriers to seeking health care in Uganda* by Nyasha Tirivayi
2014-68 *An individual-centered approach to multidimensional poverty: The cases of Chile, Colombia, Ecuador and Peru* by Andrea Franco Correa
2014-69 *Innovation and productivity in services: Empirical evidence from Latin America* by Gustavo Crespi, Ezequiel Tacsir and Fernando Vargas
2014-70 *Analysing global value chains using input-output economics: Proceed with care* by Önder Nomaler and Bart Verspagen