Structural decomposition analyses

Chen, Quanrun; Dietzenbacher, Erik; Los, Bart

Structural decomposition analyses: the differences between applying the semi-closed and the open input–output model

Quanrun Chen
School of Statistics, University of International Business and Economics, Beijing, China;
e-mail: qrchen@amss.ac.cn

Erik Dietzenbacher
Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands;
e-mail:H.W.A.Dietzenbacher@rug.nl

Bart Los
Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands;
e-mail: b.los@rug.nl

Abstract. The open and the semi-closed input–output model are widely used in impact analysis but usually yield considerably different outcomes. This paper examines whether such differences also exist in structural decomposition analyses. The empirical part considers the decomposition of the growth in gross output and in labor compensation, for China (1997–2007) and for a set of 35 other countries. Our main findings are twofold. First, for gross output growth, both models yield very similar results for the factors they have in common. Therefore, if only the contribution of these common factors is of interest, it does not matter whether the semi-closed or the open input–output model is used. Second, for the analysis of labor compensation growth, both models may yield significantly different results for the labor compensation coefficients. Therefore, if the contribution of changes in labor compensation coefficients is of interest, the semi-closed model is recommended to be used.

Keywords: input–output models, structural decomposition analysis, gross output growth, labor compensation growth

Introduction

Input–output (IO) analysis is widely used to analyze the impacts of expenditure injections at the global, the national, the regional, and even the municipal level. In the traditional IO model, all final demand categories (household consumption, government expenditures, investments, and exports) are treated as exogenous variables and determined "outside" the model. Therefore, the traditional IO model is also termed the open IO model and it only captures the linkages between industries. In many policy relevant cases, however, it is important that the linkage between industries and households is mutual. To this end household consumption is endogenized and thus determined within the model. This yields the semi-closed IO model, which has also been termed the partially closed or extended IO model (see Batey et al., 1987; Batey and Rose, 1990).

Semi-closed IO models have been widely used in policy and impact analyses. The effect of induced household consumption is usually one of the concerns and the focus is on the medium- or long-run in which case the induced household consumption can be fully effectuated. Examples at the national level include Yang et al. (2008) who evaluate the socio-economic impact of large-scale projects. Dietzenbacher and Günlük-Şenesen (2003a) assess changes in the Turkish production structure and labor income between two periods with different policy strategies. For the semi-closed model they obtain findings that cannot be detected with the open IO model, such as the dominance of public services for the industry output multipliers. The application of semi-closed IO models is even more prevalent at a

regional level. One important reason is that regional economies are more open than national economies, which reduces the importance of inter-industry linkages relative to the industry–household linkage (Batey et al., 1993; Trigg and Madden, 1994). For example, for the region of Evros in Greece, Hewings and Romanos (1981) find that 50% of the important coefficients are related to the household sector, and in an interregional IO analysis for the UK economy, McGregor et al. (1999) find that the migration effect via the income–consumption relationship is more important than the spillover and feedback trade effect. A very recent example where the semi-closed IO model has been used is the study of the impacts of higher education institutions for the Scottish economy (Hermannsson et al., 2013, 2014a, 2014b).

Next to the widespread application of semi-closed IO models, another important development in IO in the last 30 years was structural decomposition analysis (SDA). Essentially, it examines how the change over time in a variable of interest is decomposed into the effects of changes in its constituent factors. In this way the contribution of each factor to the growth of the variable of interest is obtained (see Rose and Casler, 1996, or Miller and Blair, 2009, for overviews). In the past, many applications focused on decomposing changes in economic phenomena and environmental issues. Recent examples include: Oosterhaven and Broersma (2007) who studied labor productivity; Pei et al. (2011) China's import growth; Cazcarro et al. (2013) water consumption in Spain; and Arto and Dietzenbacher (2014) global greenhouse gas emissions.

So far, however, to the best of our knowledge all SDA applications have only been for the open IO model. Schumann (1994) deems that this is because decomposing a structural change in isolated sources within the semi-closed IO model is impossible. In this study, we extend earlier work by Dietzenbacher and Günlük-Şenesen (2003b) and show that carrying out SDAs in a semi-closed IO model is very well possible.

Empirical impact studies have shown the importance of taking the industry–household linkage into full account. This immediately raises the question what the differences are between applying SDA to the semi-closed model and to the open model. In the open IO model, a change in a domestic input coefficient only causes changes in gross outputs via interindustry linkages. In the semi-closed IO model, however, these changes in gross outputs further cause changes in labor income and thus changes in household consumption, leading to additional changes in gross output. Hence, the change in gross output solely due to a change in a single domestic input coefficient will be larger (either in positive or negative terms) in the semi-closed model than in the open model. The magnitudes of these differences can only be assessed by empirical research. Furthermore, most often the effects of changes in sets of determinants (such as the matrix with all domestic input coefficients) are isolated. Typically, some elements of these sets have grown, while others have become smaller. Consequently, the effects may cancel out. One would thus expect the decomposition results from both models to be different, but the magnitude of the differences is far from clear.

Our decomposition framework is also relevant for another, further extension of the IO model. These are models based on a Social Accounting Matrix (SAM) rather than on an IO table (see Pyatt, 1988). SAM-based models are broader than semi-closed IO models as they extend the linkages to relations between institutions (firms, households, the public sector and the foreign sector). They cover production and income generation (just as IO models do), but also the distribution and re-distribution of income. Theoretically, semi-closed IO models can be derived as simplifications of SAMs by using the apportionment method (see Pyatt, 2001). The decomposition proposed in this paper can therefore also be applied to a SAM-based model (and some parts have been done in Llop, 2007).[1] We would like to point out that the choice for a semi-closed IO model rather than a SAM-based model is usually because

[1]A decomposition in a CGE framework cannot be done according to an analytical expression but requires complex numerical simulations, a rare example of which was given by Jensen-Butler and Madsen, 2005).

of a lack of data (it is easier to compile IO tables than SAMs), not because semi-closed IO models are superior to SAM-based models. On the contrary, SAM-based models are superior to semi-closed models in terms of data richness.

The remainder of this paper is organized as follows. "The open and semi-closed IO models" section introduces the open IO model and the semi-closed IO model and "Structural decomposition in the open and in the semi-closed IO model" section discusses the decompositions for the two models. "An application to Chinese IO tables" section deals with the application to the Chinese IO tables and "Findings for other countries" section takes 35 other countries into consideration. "The dependency problem: Correlation checks" section checks the plausibility of one of the underlying assumptions of the decompositions and "Conclusions" section concludes.

The open and semi-closed IO models

The traditional, open IO model is usually expressed as follows (see Miller and Blair, 2009)

$$x = (I - A)^{-1} f = L(c + g) \tag{1}$$

where x is the n-element vector of gross outputs, A the $n \times n$ matrix with domestic input coefficients, and f the n-element vector of final demands for domestic products, which can be further split into a vector of household consumption c and a vector of other final demands g (including government expenditures, investments, and exports). The matrix $L \equiv (I - A)^{-1}$ is the Leontief inverse. In equation (1), all final demand categories are treated as exogenous variables.

Production by industry i requires the input of labor and the amount of labor compensation paid is given by w_i. It includes compensation of employees and income of self-employed individuals. We have $w_i = b_i x_i$, where b_i is the labor compensation coefficient which gives the labor compensation in industry i per unit of its gross output. This yields

$$w = \hat{b}x \tag{2}$$

where w is the n-element vector with industry labor compensations and $\hat{b}$ the diagonal matrix with labor compensation coefficients.

In the open model, household consumption affects production and labor compensation. In reality, however, (part of) the labor compensation flows as labor income to households who spend it on buying the products produced by the industries. Industries and households are then mutually connected through this income–consumption relationship. The semi-closed IO model takes it into full account by endogenizing household consumption and total labor compensation (Miller and Blair, 2009). The semi-closed model can be expressed as

$$\begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} A & r\bar{c} \\ b' & 0 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} + \begin{bmatrix} g \\ 0 \end{bmatrix} \tag{3}$$

where $w = b'x = \Sigma_i w_i$ is the total labor compensation, r is the ratio between total household consumption of domestic products and total labor compensation (i.e. $r = \Sigma_i c_i / \Sigma_i w_i$), and $\bar{c}$ is the vector with consumption shares (i.e. $\bar{c}_i = c_i / \Sigma_i c_i$). Note that a change in r implicitly also covers the effects of substituting the consumption of domestic products for imported products or savings instead of consumption. Note furthermore that consumers may have other sources of income (e.g. capital income) from which they may pay for their consumption. In principle, r could thus even be larger than one.

In the traditional semi-closed model, only total labor compensation is modeled. Next, we will show that labor compensation at industry level can also be modeled in the semi-closed model, which allows for investigations of the determinants of labor compensation for each

industry. The expression for the outputs in equation (3) can be written as $x = Ax + r\bar{c}w + g$. Using $w = \Sigma_i w_i = i'w$, with i' the row summation vector consisting of ones, yields $x = Ax + r\bar{c}i'w + g$. The industry labor compensations are given in equation (2). Taken together in a single equation, this yields

$$\begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} A & r\bar{c}i' \\ \hat{b} & O \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} + \begin{bmatrix} g \\ 0 \end{bmatrix} \tag{4}$$

and the solution is given by[2]

$$\begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} I - A & -r\bar{c}i' \\ -\hat{b} & I \end{bmatrix}^{-1} \begin{bmatrix} g \\ 0 \end{bmatrix} \tag{5}$$

When compared to the open model, the semi-closed model includes extra feedback loops through the income–consumption relationship. Consider, for instance, the effects of an increase in the exports. In the open model, this will cause an increase in the gross outputs via industry linkages and an increase in the industry labor compensations. In the semi-closed model, however, the extra labor income for households will cause a further effect on the gross outputs via increased household consumption, and so forth. These extra loops are not included in the open model.

Structural decomposition in the open and in the semi-closed IO model
The aim is to decompose changes in gross output and labor compensation levels in each industry. The outcomes are the contributions to output growth and labor compensation growth of the changes in the underlying factors, one of which are the changes in the consumption shares. Because we want to compare the decomposition results across two models, it is important that both decompositions have (as much as possible) the same underlying factors. Therefore, we first reformulate both models.

Model reformulation
Equation (1), for the open model, can be rewritten as:

$$x = \lambda L\bar{f} = \lambda(I - A)^{-1}[\alpha\bar{c} + (1-\alpha)\bar{g}] \tag{6}$$

where $\lambda = \Sigma_i f_i$ is the total final demand (for domestic products); $\bar{f} = f / \lambda$ is a vector of industry shares in final demand; α is the share of total household consumption in total final demand ($\Sigma_i c_i / \lambda$) and $(1-\alpha)$ is the share of other final demand in total final demand ($\Sigma_i g_i / \lambda$); $\bar{c}$ is again the vector with consumption shares; and $\bar{g}$ is the vector of other final demand shares ($\bar{g}_i = g_i / \Sigma_i g_i$). Equation (6) distinguishes the factors that reflect the structure of the economy from a factor indicating the scale of the economy. The component $L[\alpha\bar{c} + (1-\alpha)\bar{g}]$ includes four structural factors and λ gives the scale, which has an identical effect on the gross output of each industry.

In the context of the open model, the vector w with industry labor compensations can be expressed—by combining equations (2) and (6)—as

$$w = \hat{b}x = \lambda\hat{b}(I - A)^{-1}[\alpha\bar{c} + (1-\alpha)\bar{g}] \tag{7}$$

---

[2] Equation (4) implies that $\begin{bmatrix} x \\ w \end{bmatrix} \geq \begin{bmatrix} A & r\bar{c}i' \\ \hat{b} & O \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}$ which (under mild conditions) guarantees that the inverse in equation (5) exists (see e.g. Takayama, 1985: 392).

In the semi-closed model, the gross outputs and the labor compensations are determined simultaneously. Equation (5) can be rewritten as:

$$\begin{bmatrix} x \\ w \end{bmatrix} = \mu \begin{bmatrix} I - A & -r\bar{c}i' \\ -\hat{b} & I \end{bmatrix}^{-1} \begin{bmatrix} \bar{g} \\ 0 \end{bmatrix} \tag{8}$$

where $\mu = \Sigma_i g_i$, which reflects the scale effect. The remaining part covers the structure effects.

After the model reformulation, A, $\bar{c}$, and $\bar{g}$ become the common factors in the gross output expressions of the two models, and A, $\bar{c}$, $\bar{g}$, and $\hat{b}$ become the common factors in the labor compensation expressions. Model-specific factors are λ and α for the open model and μ and r for the semi-closed model.

Next, we turn to decomposing the gross output growth and the labor compensation growth in the open model and the semi-closed model, respectively. When carrying out the decomposition, the effect of changes in scale can be nicely separated from the effects of changes in structure when a multiplicative—rather than an additive—structural decomposition analysis (MSDA, see Dietzenbacher et al., 2000, for an introduction) is used.

MSDA for the open model

We use superscripts "1" and "0" to indicate the final and initial year of a variable during a period of interest. Then, the growth of gross output during a period can be expressed as[3]

$$\frac{x^1}{x^0} = \frac{\lambda^1 L^1 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]}{\lambda^0 L^0 [\alpha^0 \bar{c}^0 + (1 - \alpha^0)\bar{g}^0]}$$

$$= \frac{\lambda^1 L^1 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]}{\lambda^0 L^0 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]} \times \frac{\lambda^0 L^1 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]}{\lambda^0 L^0 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]} \times \frac{\lambda^0 L^0 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]}{\lambda^0 L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^1]}$$

$$\times \frac{\lambda^0 L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^1]}{\lambda^0 L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^0]} \times \frac{\lambda^0 L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^0]}{\lambda^0 L^0 [\alpha^0 \bar{c}^0 + (1 - \alpha^0)\bar{g}^0]}$$

$$= \frac{\lambda^1 i}{\lambda^0 i} \tag{9.1}$$

$$\times \frac{L^1 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]}{L^0 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]} \tag{9.2}$$

$$\times \frac{L^0 [\alpha^1 \bar{c}^1 + (1 - \alpha^1)\bar{g}^1]}{L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^1]} \tag{9.3}$$

$$\times \frac{L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^1]}{L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^0]} \tag{9.4}$$

$$\times \frac{L^0 [\alpha^1 \bar{c}^0 + (1 - \alpha^1)\bar{g}^0]}{L^0 [\alpha^0 \bar{c}^0 + (1 - \alpha^0)\bar{g}^0]} \tag{9.5}$$

Equation (9.1) gives the scale effect, that is the output growth (which is the same in each industry) if the total of all final demands had increased as it actually did, while anything else

---

[3] For any two matrices (or vectors) U and V, U/V and $U \times V$ denote elementwise division and elementwise multiplication, respectively. That is, u_{ij}/v_{ij} and $u_{ij} \times v_{ij}$.

had remained constant. In a similar fashion, equations (9.2)–(9.5) give the effects of changes in $\mathbf{A}$, $\bar{c}$, $\bar{g}$, and α on the growth of gross output in each industry, respectively.[4]

It should be stressed that the decomposition results vary across different "orderings" of changing the factors. In equations (9), we first changed l, then $\mathbf{A}$, etcetera. Denote this ordering as $\lambda \to \mathbf{A} \to \bar{c} \to \bar{g} \to \alpha$. The ordering for the "mirror decomposition" is given by $\alpha \to \bar{g} \to \bar{c} \to \mathbf{A} \to \lambda$ and the decomposition itself yields

$$\frac{x^1}{x^0} = \frac{\lambda^1 \mathbf{L}^1 [\alpha^1 \bar{c}^1 + (1-\alpha^1)\bar{g}^1]}{\lambda^0 \mathbf{L}^0 [\alpha^0 \bar{c}^0 + (1-\alpha^0)\bar{g}^0]}$$

$$= \frac{\lambda^1 i}{\lambda^0 i} \tag{10.1}$$

$$\times \frac{\mathbf{L}^1 [\alpha^0 \bar{c}^0 + (1-\alpha^0)\bar{g}^0]}{\mathbf{L}^0 [\alpha^0 \bar{c}^0 + (1-\alpha^0)\bar{g}^0]} \tag{10.2}$$

$$\times \frac{\mathbf{L}^1 [\alpha^0 \bar{c}^1 + (1-\alpha^0)\bar{g}^0]}{\mathbf{L}^1 [\alpha^0 \bar{c}^0 + (1-\alpha^0)\bar{g}^0]} \tag{10.3}$$

$$\times \frac{\mathbf{L}^1 [\alpha^0 \bar{c}^1 + (1-\alpha^0)\bar{g}^1]}{\mathbf{L}^1 [\alpha^0 \bar{c}^1 + (1-\alpha^0)\bar{g}^0]} \tag{10.4}$$

$$\times \frac{\mathbf{L}^1 [\alpha^1 \bar{c}^1 + (1-\alpha^1)\bar{g}^1]}{\mathbf{L}^1 [\alpha^0 \bar{c}^1 + (1-\alpha^0)\bar{g}^1]} \tag{10.5}$$

In the case of n factors there are $n!$ equivalent decompositions (see Dietzenbacher and Los, 1998), but it turns out that the average of two "mirror" decompositions closely approximates the average of all $n!$ equivalent decompositions (de Haan, 2001). Taking the geometric average of the decompositions in equations (9) and (10) yields:

The effect of a change in λ: $E\lambda = \sqrt{(9.1) \times (10.1)} = \lambda^1 / \lambda^0$;

the effect of changes in $\mathbf{A}$: $E\mathbf{A} = \sqrt{(9.2) \times (10.2)}$;

the effect of changes in $\bar{c}$: $E\bar{c} = \sqrt{(9.3) \times (10.3)}$;

the effect of changes in $\bar{g}$: $E\bar{g} = \sqrt{(9.4) \times (10.4)}$;

the effect of a change in α: $E\alpha = \sqrt{(9.5) \times (10.5)}$.

The growth of labor compensation can be decomposed in similar way, starting from

$$\frac{w^1}{w^0} = \frac{\hat{\mathbf{b}}^1 x^1}{\hat{\mathbf{b}}^0 x^0} = \frac{\lambda^1 \hat{\mathbf{b}}^1 \mathbf{L}^1 [\alpha^1 \bar{c}^1 + (1-\alpha^1)\bar{g}^1]}{\lambda^0 \hat{\mathbf{b}}^0 \mathbf{L}^0 [\alpha^0 \bar{c}^0 + (1-\alpha^0)\bar{g}^0]}$$

---

[4] Since $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}$, the effects of $\mathbf{L}$ and $\mathbf{A}$ are equivalent.

The decomposition includes also the effects of changes in b (that is, Eb), next to the effects that are also included in the output growth decomposition. The detailed equations are given in Appendix A of the online Supplementary Material.

MSDA for the semi-closed model

In the context of a semi-closed model, the growth of gross output and labor compensation is decomposed simultaneously. The decomposition for the ordering $\mu \to \mathbf{A} \to \bar{c} \to \bar{g} \to b \to r$ yields:

$$\frac{\begin{bmatrix} x^1 \\ w^1 \end{bmatrix}}{\begin{bmatrix} x^0 \\ w^0 \end{bmatrix}} = \frac{\mu^1 \begin{bmatrix} \mathbf{I} - \mathbf{A}^1 & -r^1\bar{c}^1 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^1 \\ 0 \end{bmatrix}}{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^0\bar{c}^0 i' \\ -\hat{b}^0 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^0 \\ 0 \end{bmatrix}}$$

$$= \frac{\mu^1 i}{\mu^0 i} \tag{11.1}$$

$$\times \frac{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^1 & -r^1\bar{c}^1 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^1 \\ 0 \end{bmatrix}}{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^1 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^1 \\ 0 \end{bmatrix}} \tag{11.2}$$

$$\times \frac{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^1 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^1 \\ 0 \end{bmatrix}}{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^0 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^1 \\ 0 \end{bmatrix}} \tag{11.3}$$

$$\times \frac{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^0 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^1 \\ 0 \end{bmatrix}}{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^0 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^0 \\ 0 \end{bmatrix}} \tag{11.4}$$

$$\times \frac{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^0 i' \\ -\hat{b}^1 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^0 \\ 0 \end{bmatrix}}{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^0 i' \\ -\hat{b}^0 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^0 \\ 0 \end{bmatrix}} \tag{11.5}$$

$$\times \frac{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^1\bar{c}^0 i' \\ -\hat{b}^0 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^0 \\ 0 \end{bmatrix}}{\mu^0 \begin{bmatrix} \mathbf{I} - \mathbf{A}^0 & -r^0\bar{c}^0 i' \\ -\hat{b}^0 & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \bar{g}^0 \\ 0 \end{bmatrix}} \tag{11.6}$$

The "mirror" decomposition and the geometric averages are detailed in Appendix A of the online Supplementary Material.

This MSDA for the semi-closed model might suffer from the "dependency problem" highlighted by Dietzenbacher and Los (2000). All changes in domestic intermediate input coefficients in a column of the matrix **A** are automatically accompanied by an equally large change with opposite sign in the sum of value added and import coefficients, since column sums of coefficients always add up to one. These two sets of coefficients cannot change independently, which invalidates the main idea of SDA to quantify what would have happened if one determinant would have changed as it actually did, while other factors remained unchanged. Since the labor compensation coefficients in **b** are part of the value added coefficients, it remains an empirical question whether changes in the elements of **A** are reflected in opposite changes in **b**, or whether the other value added components and imports absorb the changes in **A**. In the latter case, the dependency problem would be irrelevant, in the former case an alternative decomposition approach introduced by Dietzenbacher and Los (2000) should be adopted. This issue will be analyzed empirically in "The dependency problem: Correlation checks" section.

An application to Chinese IO tables
Data description
We employ the MSDA to investigate the sources of the growth in industry gross output levels and in industry labor compensation levels in China during 1997–2007. We use the 1997 and 2007 survey-based Chinese IO tables as published by the National Bureau of Statistics of China (National Bureau of Statistics (NBS), 1999, 2009). These IO tables are in current prices and have 42 industries (see the industry classification in Appendix B of the online Supplementary Material). Unfortunately, IO tables in constant prices are not available. A consequence is that the effect of, for example, increased consumption on output growth not only covers quantity but also price changes. For many purposes one would want to single out the effects of quantity changes. It should be stressed though that the primary interest in this paper is the differences between decomposing the semi-closed model and the open model. Therefore, we have chosen to work with the tables in current prices.

Labor compensation in our analysis consists of two parts: compensation of employees and income of self-employed individuals. This distinction is readily available in the Chinese IO table for 1997, but not for 2007. This is because the NBS changed the statistical concepts regarding labor compensation in 2004 (see Bai and Qian, 2010). For *Agriculture* (industry 1), labor compensation in 2007 includes also the operating surplus of state-owned and collectively-owned farms. This is because obtaining detailed financial statements from these farms had become more and more difficult. For non-agricultural industries, on the other hand, the income of self-employed individuals is no longer included as part of the labor compensation in 2007. In both cases we need to make adjustments to make the 1997 and 2007 table comparable (see Appendix C of the online Supplementary Material for the adjustment details).

In addition to the adjustments for labor compensation in 2007, adaptations in both years are required to single out the domestic deliveries. That is, the IO models introduced in "The open and semi-closed IO models" section distinguish between domestic and imported products. However, the original IO tables published by NBS do not make such a distinction. The deliveries of intermediate goods and final goods need to be split according to their origin, i.e. domestically produced or imported. For this we apply the frequently used proportional approach. Multiplying each row i of the IO table with its "domestic product share in total domestic demand" φ_i, gives the domestic products delivered to industries, to households, to the government and other final demand categories. For industry i, its domestic product share

in total domestic demand is defined as $\varphi_i = (x_i - e_i)/(x_i + m_i - e_i)$, where x_i is the gross output of product i, m_i is the import of product i, and e_i is the export of product i. The import tables or matrices can be obtained by deducting the domestic products from the original IO tables.

Results and findings

The results for the decomposition of the gross output growth and the labor compensation growth during 1997–2007 are shown in Table 1 for the open model and in Table 2 for the semi-closed model. For example, output in *Agriculture* (industry 1) grew by 93% and labor compensation by 81%. That is, $x_1^{2007}/x_1^{1997} = 1.93$ and $w_1^{2007}/w_1^{1997} = 1.81$. If only the consumption shares would have changed as they did between 1997 and 2007, while anything else had remained constant, the agricultural output would have been decreased by 32% (i.e. $E\bar{c}$ equals 0.68 for industry 1) when the open model is applied. Multiplying the appropriate effects yields agricultural output growth (i.e. $1.93 = 3.82 \times 0.93 \times 0.68 \times 0.95 \times 0.84$) and multiplying all effects yields the growth in labor compensation for agriculture (i.e. $1.81 = 3.82 \times 0.93 \times 0.68 \times 0.95 \times 0.84 \times 0.94$).

The results for the three factors in the decompositions of output growth that both models have in common (i.e. $E\mathbf{A}$, $E\bar{c}$ and $E\bar{g}$) look fairly similar. To quantify this comparison across models we have calculated the absolute relative difference (ARD).[5] For example, for the case of $E\mathbf{A}$ it is defined as

$$\rho_i^{E\mathbf{A}} = \frac{|(E\mathbf{A})_i^{\text{semi}} - (E\mathbf{A})_i^{\text{open}}|}{\sqrt{(E\mathbf{A})_i^{\text{semi}}(E\mathbf{A})_i^{\text{open}}}} \times 100\%$$

where $(E\mathbf{A})_i^{\text{semi}}$ gives the effect of changes in the input coefficients matrix $\mathbf{A}$ on the output and labor compensation in industry i when the semi-closed model is used and $(E\mathbf{A})_i^{\text{open}}$ indicates the effect when the open model is used. It gives the ARD between $(E\mathbf{A})_i^{\text{semi}}$ and $(E\mathbf{A})_i^{\text{open}}$, and the benchmark is their geometric average. The definition of $\rho_i^{E\mathbf{A}}$ is intuitive and has several nice properties (such as satisfying the condition of symmetry, i.e. exchanging $(E\mathbf{A})_i^{\text{semi}}$ by $(E\mathbf{A})_i^{\text{open}}$ and vice versa does not change the value of $\rho_i^{E\mathbf{A}}$, and reporting $\rho_i^{E\mathbf{A}} = 0$ when the result is the same in both models).

The averages over all industries of the three ARDs are very small. That is, 0.33% for $E\mathbf{A}$, 2.69% for $E\bar{c}$, and 2.04% for $E\bar{g}$. Also their standard deviations are small (0.16%, 1.52%, and 1.19%, respectively). This outcome—which may not be very exciting in itself—bears a considerable amount of relevance. It states that the contribution to output and labor compensation growth of, for example, changes in consumption shares is more or less the same, irrespective of the model that is used. At first sight, this may come as a surprise, because the semi-closed model includes an extra channel via the industry-household linkage. Apparently this extra linkage does not play a role in decompositions.[6] Closer inspection learns that it is a matter of counterbalancing. Consider the changes in the vector with industry shares of other final demands (i.e. $\bar{g}$), which is exogenous in both models. Because it is a vector of shares, $\Delta\bar{g}$ will include pluses and minuses. The pluses will increase outputs and labor compensations, and the minuses will have an opposite effect. The effects will thus cancel each other out to a large extent and will result in a relatively small change in the total

---

[5] The results for $E\bar{c}$ and $E\bar{g}$ are always larger in Table 1 than in Table 2, which implies that it is not necessary to use absolute values in the calculation of the relative differences. This systematic difference does not hold for the results for $E\mathbf{A}$. Unlike for other industries, the result for *Construction* (industry 26) in Table 1 is smaller than in Table 2 (although the same outcome is reported due to rounding). Therefore we calculate ARDs.

[6] It should be stressed that it does play a significant role for impact analyses, for example. In decompositions, however, one is actually interested in relative aspects. A given output growth is "distributed" over the contributing factors.

Table 1. Decomposition results for gross output growth and labor compensation growth (open model).

Industry	Gross output growth	Labor compensation growth	EA	$E\bar{c}$	$E\bar{g}$	$E\alpha$	Eb
1	1.93	1.81	0.93	0.68	0.95	0.84	0.94
2	4.33	2.93	1.11	1.03	0.96	1.03	0.68
3	5.84	5.71	1.56	1.04	0.90	1.04	0.98
4	5.15	4.07	1.06	1.01	1.14	1.11	0.79
5	2.61	2.12	0.77	0.97	0.83	1.11	0.81
6	3.03	3.11	1.19	0.87	0.93	0.82	1.03
7	2.72	1.82	0.89	0.93	0.83	1.04	0.67
8	2.97	1.86	1.10	0.97	0.76	0.96	0.63
9	4.91	3.80	1.14	0.94	1.13	1.06	0.77
10	3.38	1.86	0.98	1.00	0.88	1.03	0.55
11	6.80	7.64	1.70	1.06	0.95	1.04	1.12
12	4.08	2.59	1.09	0.97	1.00	1.01	0.63
13	2.59	1.91	0.75	0.96	0.83	1.13	0.74
14	7.86	4.17	1.59	1.01	1.15	1.11	0.53
15	3.55	2.31	0.89	0.98	0.97	1.10	0.65
16	4.80	2.97	1.02	1.01	1.08	1.13	0.62
17	6.21	4.50	1.14	1.07	1.24	1.07	0.72
18	4.88	2.59	0.97	0.99	1.25	1.12	0.53
19	8.41	4.75	1.02	1.01	1.92	1.12	0.57
20	5.88	3.28	0.92	1.02	1.44	1.13	0.56
21	2.68	2.19	0.67	1.04	1.00	1.00	0.82
22	8.17	–	1.80	1.00	1.09	1.09	–
23	8.06	5.01	2.01	1.09	0.96	0.99	0.62
24	7.94	3.28	1.78	1.23	1.09	0.88	0.41
25	3.08	3.44	0.81	1.14	0.94	0.93	1.12
26	3.61	2.20	0.96	1.02	0.81	1.18	0.61

Table 1. (continued)

Industry	Gross output growth	Labor compensation growth	EA	E$\bar{c}$	E$\bar{g}$	Eα	Eb
27	6.26	3.80	1.39	1.04	1.11	1.02	0.61
28	3.65	4.36	1.43	0.71	1.00	0.94	1.19
29	5.70	6.74	0.92	1.43	1.15	0.99	1.18
30	2.61	2.91	0.68	1.09	0.95	0.97	1.12
31	4.88	4.09	1.10	1.30	0.98	0.91	0.84
32	5.42	5.31	1.24	1.17	1.05	0.94	0.98
33	7.96	6.04	1.13	1.70	1.27	0.85	0.76
34	5.33	3.97	1.02	1.10	1.23	1.01	0.75
35	4.97	4.29	1.61	1.01	0.71	1.13	0.86
36	5.76	4.39	1.83	0.99	0.75	1.11	0.76
37	1.49	1.13	0.78	0.98	0.48	1.07	0.75
38	4.38	3.50	0.85	1.48	1.01	0.89	0.80
39	5.68	4.89	1.03	1.37	1.04	1.01	0.86
40	6.19	5.07	1.09	1.37	1.16	0.94	0.82
41	3.51	2.81	1.18	0.94	0.84	0.98	0.80
42	3.57	4.54	1.01	1.00	0.77	1.20	1.27
Scale effect, Eλ		3.82					

Notes: (1) The gross output growth and labor compensation growth are measured as the ratio between their levels in 2007 and 1997; (2) The labor compensation of industry 22 in 1997 is 0, so we do not report the result for industry 22; (3) EA, $E\bar{c}$, $E\bar{g}$, $E\alpha$, and $E\lambda$ denote the effect of changes in input coefficients, consumption shares, other final demand shares, the share of total household consumption in total final demand, and total final demand on gross output growth and labor compensation growth, respectively; Eb denotes the effect of changes in labor compensation coefficients on labor compensation growth.

Table 2. Decomposition results for gross output growth and labor compensation growth (semi-closed model).

Industry	Gross output growth	Labor compensation growth	EA	E$\bar{c}$	E$\bar{g}$	Er	E$b_{(x)}$	Eb
1	1.93	1.81	0.94	0.64	0.90	0.97	0.80	0.75
2	4.33	2.93	1.12	1.01	0.94	0.98	0.91	0.61
3	5.84	5.71	1.57	1.02	0.89	0.99	0.92	0.89
4	5.15	4.07	1.06	0.99	1.13	0.99	0.95	0.75
5	2.61	2.12	0.77	0.96	0.82	0.99	0.95	0.77
6	3.03	3.11	1.20	0.82	0.89	0.96	0.79	0.81
7	2.72	1.82	0.89	0.90	0.81	0.99	0.92	0.61
8	2.97	1.86	1.11	0.94	0.74	0.98	0.87	0.54
9	4.91	3.80	1.14	0.92	1.12	0.99	0.92	0.72
10	3.38	1.86	0.99	0.97	0.86	0.98	0.91	0.50
11	6.80	7.64	1.70	1.03	0.94	0.99	0.91	1.03
12	4.08	2.59	1.09	0.95	0.98	0.98	0.90	0.57
13	2.59	1.91	0.75	0.95	0.82	0.99	0.96	0.71
14	7.86	4.17	1.59	1.00	1.14	0.99	0.95	0.51
15	3.55	2.31	0.89	0.97	0.95	0.99	0.95	0.61
16	4.80	2.97	1.02	1.00	1.07	0.99	0.96	0.60
17	6.21	4.50	1.14	1.05	1.23	0.99	0.93	0.68
18	4.88	2.59	0.97	0.97	1.23	0.99	0.93	0.49
19	8.41	4.75	1.02	0.99	1.90	0.99	0.96	0.54
20	5.88	3.28	0.92	1.01	1.43	0.99	0.97	0.54
21	2.68	2.19	0.68	1.01	0.97	0.98	0.89	0.73
22	8.17	–	1.81	0.98	1.08	0.99	0.94	–
23	8.06	5.01	2.02	1.06	0.94	0.98	0.89	0.55
24	7.94	3.28	1.79	1.17	1.04	0.97	0.82	0.34
25	3.08	3.44	0.81	1.09	0.91	0.98	0.85	0.95
26	3.61	2.20	0.96	1.02	0.81	1.00	0.99	0.61

Continued

Table 2. (continued)

Industry	Gross output growth	Labor compensation growth	EA	E$\bar{c}$	E$\bar{g}$	Er	E$b_{(x)}$	Eb
27	6.26	3.80	1.40	1.01	1.09	0.98	0.91	0.55
28	3.65	4.36	1.44	0.68	0.97	0.97	0.86	1.02
29	5.70	6.74	0.92	1.38	1.12	0.98	0.88	1.05
30	2.61	2.91	0.68	1.05	0.93	0.98	0.87	0.98
31	4.88	4.09	1.11	1.24	0.95	0.97	0.84	0.71
32	5.42	5.31	1.24	1.12	1.02	0.98	0.86	0.84
33	7.96	6.04	1.14	1.62	1.21	0.97	0.80	0.61
34	5.33	3.97	1.03	1.06	1.20	0.98	0.90	0.67
35	4.97	4.29	1.61	1.00	0.70	0.99	0.96	0.83
36	5.76	4.39	1.84	0.97	0.74	0.99	0.95	0.73
37	1.49	1.13	0.78	0.97	0.47	0.99	0.93	0.70
38	4.38	3.50	0.86	1.41	0.98	0.97	0.83	0.66
39	5.68	4.89	1.03	1.34	1.01	0.98	0.90	0.77
40	6.19	5.07	1.09	1.32	1.12	0.98	0.86	0.70
41	3.51	2.81	1.19	0.91	0.82	0.98	0.88	0.71
42	3.57	4.54	1.01	1.00	0.77	1.00	1.00	1.27

Scale effect, $E\mu$ 4.58

Notes: (1) and (2) see Table 1; (3) for E**A**, $E\bar{c}$, and $E\bar{g}$, see Note (3) of Table 1; Er and $E\mu$ denote the effects (on gross output growth and labor compensation growth) of changes in the ratio between total household consumption of domestic products and total labor compensation, and of changes in the total other final demands, respectively; because the changes in labor compensation coefficients affect the industry gross output and labor compensation differently, we use $Eb_{(x)}$ and Eb to denote the effect of changes in labor compensation coefficients on the gross output growth and on the labor compensation growth, respectively.

labor compensation. This is precisely what is fed back to households in the semi-closed model and which affects consumption, production, total labor compensation and so forth. Apparently, the feedback effects are very small in the present case, which explains why the results for both models are so similar. The same reasoning applies to the changes in $\overline{c}$ and to the changes in **A** (which usually also include increases as well as decreases).

For the decomposition of labor compensation growth, $E\mathbf{A}$, $E\overline{c}$, $E\overline{g}$, and Eb are the factors that the open model and the semi-closed model have in common. As equations (2) and (4) show, the industry labor compensation can be expressed as $\mathbf{w} = \hat{\mathbf{b}}\mathbf{x}$ both in the open model and in the semi-closed model. This means that the changes in **A**, $\overline{c}$, and $\overline{g}$ affect $\mathbf{w}$ only through their effects on the gross outputs. The effects of $E\mathbf{A}$, $E\overline{c}$, and $E\overline{g}$ on industry labor compensation growth are therefore exactly the same as their effects on gross output growth, which holds for both models. Whereas the effects of $E\mathbf{A}$, $E\overline{c}$, and $E\overline{g}$ are very similar in the two models, the effects of Eb on labor compensation growth differ significantly. The semi-closed model indicates a stronger negative (or weaker positive) effect than the open model does. The mean of the ARDs across all industries is 10.31% with a standard deviation of 5.94%. For some industries, the signs of the effects are even different. For instance, the open model indicates a small positive effect on labor compensation in *Manufacture of food products and tobacco processing* (industry 6), whereas the semi-closed model indicates a substantial negative effect.

The difference between the results for changes in labor compensation coefficients in both models is caused by the feedback loop in the semi-closed model. In the case of $E\mathbf{A}$, $E\overline{c}$, and $E\overline{g}$, we had pluses and minuses that counterbalanced so that there was no role for the feedback loop. In the case of Eb, however, we essentially have minuses (probably caused by the prominent role China started to play as a location of stages of increasingly fragmented production processes) that are enlarged via feedback effects. In the open model, the effect of Eb on labor compensation is direct and results in decreases for most industries (as shown in the last column of Table 1). All these decreases imply that total labor compensation is considerably lowered. In the semi-closed model, this reduces consumption through the industry-household linkage, which in its turn lessens output, labor compensation, and so forth. The reductions in industry labor compensations due to changes in b that we find in Table 1 for the open model are thus strengthened by the feedback loop in the semi-closed model. For no industry we find a larger value for Eb in Table 2 (for the semi-closed model) than in Table 1 (for the open model), and in most cases the value is substantially smaller.

Another difference between the decompositions in the two models is the scale effect. For the open model we have $E\lambda = 3.82$ and for the semi-closed model we find $E\mu = 4.58$. Note that the scale indicators in the two models have been defined differently, as a consequence of which the change in scale will generally also be different. Recall that $\lambda = \Sigma_i f_i = \Sigma_i (c_i + g_i)$ is total final demands, which is usually much larger (but never smaller) than $\mu = \Sigma_i g_i$, i.e. total final demands less total household consumption. The two effects will be exactly the same, only if total household consumption ($\Sigma_i c_i$) grows with the same rate as total other final demands ($\Sigma_i g_i$) does. In China, total household consumption grew by 169% in nominal terms (from 3428 billion Yuan in 1997 to 9219 billion Yuan in 2007), total other final demands grew by 358% (from 5126 billion to 23,485 billion Yuan), so that total final demands grew by 282% (from 8554 billion to 32,704 billion Yuan). The findings ($E\lambda = 3.82$ and $E\mu = 4.58$) are in line with the common viewpoint that China's output growth was primarily export-led. Exports are a major part of other final demands and grew by 478% (from 1654 billion to 9554 billion Yuan). Investment demand also grew at a much faster pace than household consumption, at a rate of 281% (from 2701 billion Yuan in 1997 to 10,293 billion Yuan in 2007), reflecting the transformation of a mainly agriculture-based to a manufacturing-based economy.

Changes in labor compensation coefficients b do not (i.e. $Eb_{(x)}$ in Table 2) play a role in the decomposition of output growth in the open model, but they do play a role in the semi-closed model. Their effects are listed in Table 2 in the column $Eb_{(x)}$. It appears that changes in b have a lowering effect on the gross output growth in each industry.[7] For industries such as *Agriculture* (industry 1), *Manufacture of food products and tobacco processing* (6) and *Real estate* (33), the effects are rather significant (decreasing output by 20% or more). For one-third of the industries the changes in labor compensation coefficients alone would have caused output to fall by more than 10%, which is quite much given the fact that the effect is entirely indirect. The labor compensation coefficients decreased during 1997–2007 for 34 (out of 42) industries. They were insufficiently counterbalanced by increased coefficients so that total labor compensation fell. This implied further decreases in the household consumption for each industry, the output, the labor compensation, and so forth. All these indirect effects add up and have in the end a considerable effect. Industries with larger consumption shares, such as *Agriculture* (industry 1), are usually affected more than industries with smaller consumption shares.

Finally, we would like to point out another difference between the two models. Unlike the open model, household consumption growth (at the industry level) is determined endogenously in the semi-closed model. Therefore consumption growth can be decomposed in the semi-closed IO model, but not in the open model. It should be stressed though that the changes in many determinants have an identical effect on the consumption growth of products from different industries. The only factor that affects consumption differently across industries is the (exogenous) consumption shares, which makes this decomposition of lesser interest. The decomposition analysis of the growth in China's household consumption is presented in Appendix D of the online Supplementary Material.

Findings for other countries

The MSDA application for 1997 and 2007 Chinese IO tables shows that the semi-closed model and the open model yield very similar results for the common factors when decomposing output growth. For the decomposition of labor compensation growth, however, both models yield significantly different results for the contribution of changes in labor compensation coefficients. In this section, the MSDA has been applied to the IO tables of a set of other countries to investigate the general validity of our findings.

A test version of the world input–output database (WIOD) was used as our data source.[8] WIOD provides world IO tables as well as socio-economic accounts on an annual basis for the period 1995–2011 (see the construction details in Dietzenbacher et al., 2013). The world IO tables are inter-country IO tables with 35 industries and 40 countries and the rest of the world (as a 41st country). Since our analysis is based on national IO tables, the 1997 and 2007 world IO tables in current prices have been aggregated, yielding the national IO tables for 40 countries. The data on labor compensation are obtained from the socio-economic accounts in WIOD, which are in national currencies. They have been converted to US dollars using the exchange rates that have been used in WIOD for the IO tables. For Indonesia, India, Turkey and Taiwan, data on labor compensation are not available and the compensation of employees has been used as an alternative.

The labor compensation growth was decomposed for the open and the semi-closed model, and Table 3 focuses on the four common factors (A, $\bar{c}$, $\bar{g}$, and b). (Note that the results for A, $\bar{c}$ and $\bar{g}$ are exactly the same for the decomposition of gross output growth.) For each country

---

[7] The output decreases also in *Public management and social administration* (industry 42), although very little. Because of rounding, Table 2 reports 1.00 for this industry.

[8] The most recent WIOD data are available free of charge at www.wiod.org.

Table 3. Decomposition results for other countries (1997–2007).

Countries	EA		E$\bar{c}$		E$\bar{g}$		Eb		γ	corr
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
Australia	0.34	0.15	0.38	0.17	1.56	0.68	1.25	0.54	−1.95	−0.41
Austria	0.37	0.29	0.09	0.07	1.55	1.23	6.84	5.41	−14.78	−0.19
Belgium	0.18	0.19	0.38	0.39	0.46	0.48	2.27	2.36	−7.60	−0.26
Bulgaria										0.06
Brazil	0.48	0.19	2.81	1.09	2.01	0.81	5.81	2.29	−5.88	0.18
Canada	2.93	1.90	0.92	0.58	0.10	0.15	3.09	2.01	−5.61	0.26
China	0.82	0.48	1.76	1.06	2.03	1.19	11.70	6.99	−30.51	−0.27
Cyprus										−0.24
Czech Republic	2.24	1.60	0.27	0.19	3.62	2.57	2.20	1.57	−5.93	0.28
Denmark	0.74	0.59	0.03	0.04	2.25	1.83	1.73	1.41	−4.58	0.36
Estonia	0.86	0.72	0.42	0.34	0.14	0.12	0.04	0.04	0.25	−0.06
Finland	0.21	0.15	0.06	0.04	0.80	0.57	5.21	3.70	−11.56	−0.22
France	1.13	0.71	0.03	0.03	0.81	0.51	3.74	2.35	−6.02	−0.11
Germany	0.96	0.59	0.10	0.07	1.93	1.21	7.43	4.67	−12.55	−0.31
Greece										0.05
Hungary	2.12	1.41	0.09	0.06	3.35	2.25	0.37	0.25	−0.78	−0.17
India	4.30	1.97	0.39	0.19	3.73	1.71	12.53	5.84	−17.20	−0.67
Indonesia	1.99	1.04	0.66	0.44	3.27	1.82	1.92	1.10	−2.94	−0.20
Ireland	1.47	1.14	0.61	0.47	1.72	1.29	3.93	3.04	−12.68	−0.48
Italy	1.89	0.94	1.13	0.56	0.82	0.41	7.31	3.67	−9.68	−0.10
Japan	0.30	0.16	1.67	0.92	1.01	0.56	3.68	2.02	−4.74	−0.15
Korea	1.47	0.96	0.65	0.43	1.41	0.93	12.23	7.99	−18.93	−0.17
Latvia	2.73	2.05	0.55	0.44	0.60	0.46	6.70	5.17	12.81	−0.32
Lithuania	0.49	0.42	0.51	0.56	0.95	0.85	2.58	2.43	3.80	−0.14
Luxembourg	0.50	0.70	0.11	0.15	2.05	3.08	5.04	7.29	−31.07	−0.27
Malta	3.11	1.94	0.51	0.36	1.25	0.97	6.83	4.41	−17.24	−0.11

Table 3. (continued)

Countries	EA		$E\bar{c}$		$E\bar{g}$		Eb		γ	corr
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
Mexico	0.56	0.31	1.08	0.62	3.51	1.99	6.71	3.85	−8.66	0.24
Netherlands	0.19	0.19	0.03	0.03	0.33	0.32	1.48	1.42	−4.73	−0.14
Poland	3.89	2.21	2.56	1.62	2.69	1.52	21.39	11.54	−27.17	0.01
Portugal	1.53	0.97	1.33	0.83	0.44	0.29	0.59	0.36	0.41	−0.41
Romania										−0.03
Russia	2.51	1.24	0.11	0.07	0.79	0.40	3.75	1.85	−6.00	−0.63
Slovak Republic	4.47	3.39	0.47	0.34	2.50	1.97	0.16	0.16	−0.88	0.07
Slovenia	2.79	2.25	0.27	0.23	1.39	1.15	3.85	3.20	−8.67	0.09
Spain	0.06	0.03	0.38	0.19	0.28	0.14	9.74	4.74	−14.36	−0.10
Sweden	0.62	0.48	0.68	0.53	0.54	0.42	2.86	2.23	−7.26	0.05
Taiwan	0.85	0.64	1.16	0.88	5.42	4.02	9.37	7.00	−15.35	−0.75
Turkey	7.81	3.21	4.58	1.87	12.20	5.09	19.31	8.04	−20.91	−0.35
United Kingdom	0.11	0.06	0.04	0.04	0.56	0.34	1.29	0.71	1.11	−0.50
USA	0.89	0.39	0.06	0.03	0.18	0.08	1.00	0.43	−0.80	−0.72

Notes: (1) For EA, $E\bar{c}$, $E\bar{g}$, and Eb, see Note (3) of Table (2); *corr* denotes the correlation coefficient between the change in labor compensation coefficients; *corr* denotes the correlation coefficient between the change in the column sum of the domestic input coefficient matrix and the change in the corresponding labor compensation coefficient during 1997–2007; (2) Mean and SD denote the mean and standard deviation of the ARDs (%) across all industries, respectively.

and for each common factor, it gives the mean and standard deviation of the ARDs using all 35 industries. Note that results for Bulgaria, Cyprus, and Greece have not been included. This is because for some industries other final demands show negative entries (because of negative changes in inventories) which leads to problems in a multiplicative decomposition. Also Romania is left out because the labor compensation of agriculture is in 2007 larger than its gross output. This is caused by a large negative compensation of capital in Romanian agriculture in that year.

Table 3 shows that, in general, the open and the semi-closed model yield quite similar decomposition results for their common factors $\mathbf{A}$, $\overline{c}$, and $\overline{g}$. For almost all countries, the means of the ARDs are less than 5% and show quite small standard deviations. For more than half of the countries, the means are even less than 1%. Turkey is the only exception where the two models yield significantly different results for the factor $\overline{g}$ (i.e. other final demand shares). The mean of the ARDs is 12.20%. The reason is that the changes in $\overline{g}$ caused a substantial increase in the gross output of industries with very high labor compensation coefficients, such as *Public administration and defence, including compulsory social security* and *Education*. This led to a clear increase in total labor compensation and as a consequence the feedback loop started to play a role. According to the Turkish IO tables, the other final demand shares of these two industries increased from 7.33% and 0.34% in 1997 to 13.13% and 5.77% in 2007, which was due to a substantial increase in government expenditures (which are part of other final demands). The average of the 1997 and 2007 labor compensation coefficients of these industries are 0.76 and 0.49, which is much larger than those in other industries.[9] This change in $\overline{g}$ thus causes a substantial increase in total labor compensation, which is further transmitted in the semi-closed model resulting in additional increases in gross outputs. The same applies (albeit to a lesser extent) for the changes in $\mathbf{A}$ in Turkey, which yields a considerable difference between the models (the mean of the ARDs being 7.81%).

For the effects of changes in the labor compensation coefficients $\mathbf{b}$, the magnitude of the difference between the two models varies across countries. For China, India, Korea, Poland, and Turkey, the means of ARDs exceed 10% and also the standard deviations are relatively large. For Austria, Brazil, Finland, Germany, Spain, Italy, Latvia, Luxembourg, Mexico, Malta, and Taiwan, the means of ARDs are all larger than 5% and the standard deviations are considerable. However, for other countries, especially Estonia, Hungary, Portugal, Slovak Republic, and USA, the ARDs are relatively small.

To analyze the role of changes in the labor compensation coefficients in the differences between the two model outcomes, we compute the average percentage change in labor compensation coefficients, weighted by the average gross output levels of the industries in the 1997 and 2007:

$$\gamma = \frac{(\mathbf{b}_1 - \mathbf{b}_0)'(\mathbf{x}_0 + \mathbf{x}_1)/2}{(w_0 + w_1)/2} \times 100\% = \frac{(\mathbf{b}_1 - \mathbf{b}_0)'(\mathbf{x}_0 + \mathbf{x}_1)}{(w_0 + w_1)} \times 100\% \tag{12}$$

where $\mathbf{b}_0$ and $\mathbf{b}_1$ are the labor compensation coefficient vectors for the initial and final year, $\mathbf{x}_0$ and $\mathbf{x}_1$ the gross output vectors for the initial and final year, w_0 and w_1 the total labor compensations for the initial and final year, respectively. It can be seen from the results listed in the tenth column of Table 3 that countries (such as China, India, Korea, Poland and Turkey) with large (weighted) average changes in labor compensation coefficients show very significant differences in $E\mathbf{b}$ between the two models. Oppositely, countries with smaller average changes in labor compensation coefficients (such as Estonia, Hungary, Portugal,

---

[9] The average labor compensation coefficients of other sectors are 0.16 in 1997 and 0.14 in 2007.

Slovak Republic and USA) show very small differences in *Eb* as obtained by the two models. Therefore, when decomposing labor compensation growth, we suggest to check first whether the changes in the labor compensation coefficients of large industries are substantial and generally have similar signs. If not, it does not matter whether the open model or the semi-closed model is used. If these changes are substantial and have often the same sign, the semi-closed model is recommended to be used, since it is more realistic than the open model by incorporating the income–consumption relationship.

The analysis of the sources of change in labor compensation over the period 1997–2007 for the countries included in WIOD suggests that the results obtained in "An application to Chinese IO tables" section for the specific case of China generally carry over to other countries with rapidly decreasing labor compensation coefficients. For such countries, we find that the contribution of changes in these coefficients lead to effects that depend quite a bit on the type of model used. For countries with labor compensation coefficients that are less dynamic than those in China, the choice of underlying model matters much less. Generally speaking, we find that the changes in domestic intermediate input coefficients, the changes in industry shares in consumption and changes in industry shares in other final demand cause rather similar changes in labor compensation levels for both models. This does not only apply to China, but to almost all countries analyzed in this section.

The dependency problem: Correlation checks

SDA disentangles the change in a variable (e.g. output) into the contributions of changes in each of its determinants (e.g. household consumption). The typical result in this example expresses what the output change would have been if only consumption had changed as it actually did while anything else had remained constant. Dietzenbacher and Los (2000) drew attention to the fact that this involves an implicit assumption, namely that all determinants are independent. The plausibility of this assumption is checked in this section.

In IO models, domestic input coefficients, import coefficients, the labor compensation coefficient, and other value added coefficients sum columnwise to 1, which holds for each industry. In the present case, the domestic input coefficient matrix $\mathbf{A}$ and the labor compensation coefficients in $\mathbf{b}$ may not be strictly independent due to this adding-up constraint. Dietzenbacher and Los (2000) show that dependencies may cause a bias in the results of decomposition analyses and how this can be remedied. The magnitude of the dependency between the column sums of $\mathbf{A}$ and the corresponding elements in $\mathbf{b}$ is uncertain. For instance, a change in the column sums of $\mathbf{A}$ is not necessarily related to a change in the labor compensation coefficient, it may be absorbed by any of the other factors (import coefficients and/or other value added coefficients). Therefore, we check the correlation between the column sums of $\mathbf{A}$ and $\mathbf{b}$ in practice. If the correlation is not strong, then the dependency will not affect the accuracy of the decompositions.

The correlation is first checked for China based on the 1997 and 2007 Chinese IO tables obtained from NBS (1999, 2009). We have calculated for each industry j the change between 1997 and 2007 in: the jth column sum of $\mathbf{A}$ (i.e. $\Sigma_i \Delta a_{ij}$); the change in the jth labor compensation coefficient (i.e. Δb_j); the change in the jth import coefficient (i.e. $\Delta \omega_j$, where $\omega_j = m_j / x_j$); and the jth coefficient for other value added items per unit of gross output (i.e. Δv_j). Next we have calculated the correlation coefficient for any pair of variables, although our prime interest is in the correlation between $\Sigma_i \Delta a_{ij}$ and Δb_j. The results are given in Table 4.

We observe a quite strong negative correlation (−0.73) between the column sums of the domestic input coefficient matrix and the other value added coefficients. This implies that a decrease in the domestic inputs share of an industry is often counterbalanced by an increase in this industry's coefficient for other value added items. This coefficient does not play a role in our decomposition and will not affect our results. The correlations between the changes

Table 4. The correlation coefficients between the changes in coefficients (1997–2007).

	$\Sigma_i \Delta a_{ij}$	$\Delta \omega_j$	Δb_j	Δv_j
$\Sigma_i \Delta a_{ij}$	1			
$\Delta \omega_j$	−0.25	1		
Δb_j	−0.22	−0.28	1	
Δv_j	−0.73	−0.12	−0.29	1

Note: $\Sigma_i \Delta a_{ij}$ denotes the change in the jth column sum of the domestic input coefficient matrix; $\Delta \omega_j$ denotes the change in the jth import coefficient; Δb_j denotes the change in the jth labor compensation coefficient; and Δv_j denotes the change in the jth coefficient for other value added items per unit of gross output.

in other coefficients are all quite small. In particular, the one correlation that might have an effect on our results (i.e. between the column sums of the domestic input coefficient matrix and the labor compensation coefficients) is only −0.22.[10] This implies that dependency between determinants is not an issue in our decompositions and will not affect their accuracy. We can therefore carry out the decompositions in "Structural decomposition in the open and in the semi-closed IO model" section without any adaptations (as suggested in Dietzenbacher and Los, 2000, for the case when dependencies are present).

The correlation between the column sums of the domestic input coefficient matrix and the corresponding labor compensation coefficients has also been checked for other countries based on the national IO tables obtained from the WIOD. The results in the last column of Table 3 indicate that the correlation is quite weak for most countries. Therefore, the accuracy of decomposition for these countries will not be affected by the dependency issue. However, we also find that some countries such as India, Russia, Taiwan, and USA show strong negative correlations. This implies that for these countries changes in the domestic inputs share of an industry is often counterbalanced by changes in this industry's labor compensation coefficient. Most probably, this is due to the fact that most of these countries are big economies that do not depend very much on imports. Small import coefficients cannot absorb large changes in the domestic input coefficients, as a consequence of which the labor compensation coefficients are likely to absorb more of these changes. In this situation, the accuracy of decomposition for these countries will be affected by the strong correlations. If the decomposition for these countries is of interest, the dependency issue is suggested to be tackled following Dietzenbacher and Los (2000).

Conclusions

The standard open IO model takes household consumption exogenous, whereas it is endogenized in the semi-closed IO model. Production leads to labor compensation which is then fed back to households which leads to consumption. This industry–households linkage provides an extra channel in the semi-closed model when compared to the open model. In impact analyses, for example, the outcomes of the two models have been found to be rather different. The question we have addressed in this paper is whether this also applies to structural decomposition analyses (to determine how much of the change in an endogenous variable is due to the change in each of the constituent exogenous factors). To this end we

---

[10] This correlation coefficient is statistically insignificant at a 5% significance level.

have decomposed the growth between 1997 and 2007 in industry output levels and in labor compensation levels for China and for 35 other countries, using both models.

First, we found for the decomposition of gross output growth in China that the open model and the semi-closed model yielded very similar results for the factors that the models have in common. These are: the changes in the domestic input coefficients matrix; the changes in the vector with consumption shares; and the vector with shares for other final demands (including government expenditures, exports, and investments). The decomposition analyses for 35 other countries showed this finding to hold in general. Hence, when decomposing gross output growth, for the common factors it does not matter whether the semi-closed model or the open model is used. A difference between the two models is that in the semi-closed model also the changes in labor compensation coefficients contribute to gross output growth (through the feedback effects that are not present in the open model). This contribution turned out to be quite significant in the study for China.

Second, in the decomposition of labor compensation growth, the magnitude of the difference between two models regarding the effects of changes in labor compensation varied across countries. When the direct effect of changes in labor compensation coefficients on total labor compensation was sizable, both models yielded significantly different results, otherwise the differences were very small. The reason is that the direct effect on total labor compensation (which is also captured by the open model) is further enlarged by the feedback loops in the semi-closed model. Then, the level of this direct effect plays an important role in determining the difference between the two models. Hence, for the decomposition analysis of labor compensation growth, a simple check on the direct effect of changes in labor compensation coefficients on total labor compensation is suggested to be done first. If this direct effect is sizable, then the semi-closed model is recommended to be used since it is more realistic than the open model by incorporating the income–consumption relationship.

In summary, these findings give some implications for model selections between the open model and the semi-closed model in the practice of SDA. For the decomposition of gross output or gross output related variables (such as CO_2 emissions, imports), if only the contribution of common factors (i.e. the domestic input coefficient matrix, the consumption share vector, the other final demand share vector) in both models is of interest, it does not matter whether the open model or the semi-closed model is used. For the decomposition of labor compensation, if the contribution of changes in labor compensation coefficients is of interest and its direct effect on total labor compensation is sizable, then an SDA of the semi-closed model is recommended. This is because on the one hand the semi-closed is more realistic than the open model by incorporating the income–consumption relationship and on the other hand because an SDA of the open model will yield significantly different results. However, we should notice that in some cases the semi-closed model based SDA could suffer from serious dependency issues. In this situation, further treatments as suggested by Dietzenbacher and Los (2000) are required to tackle the dependency issue.

Supplementary material
Appendices A–D

References

Arto I, Dietzenbacher E, 2014, "Drivers of the growth in global greenhouse gas emissions" *Environmental Science & Technology* **48**(10) 5388–5394

Bai C, Qian Z, 2010, "The factor income distribution in China: 1978-2007" *China Economic Review* **21**(4) 650–670

Batey P, Madden M, Scholefield G, 1993, "Social-economic impact assessment of large-scale projects using input-output analysis: a case study of an airport" *Regional Studies* **27**(3) 179–191

Batey P, Madden M, Weeks M J, 1987, "Household income and expenditure in extended input-output models: a comparative theoretical and empirical analysis" *Journal of Regional Science* **27**(3) 341–356

Batey P, Rose A Z, 1990, "Extended input-output models: progress and potential" *International Regional Science Review* **13**(1&2) 27–49

Cazcarro I, Duarte R, Sánchez-Chóliz J, 2013, "Economic growth and the evolution of water consumption in Spain: a structural decomposition analysis" *Ecological Economics* **96** 51–61

de Haan M, 2001, "A structural decomposition analysis of pollution in the Netherlands" *Economic Systems Research* **13**(2) 181–196

Dietzenbacher E, Günlük-Şenesen G, 2003a, "Demand-pull and cost-push effects on labor income in Turkey, 1973-90" *Environment and Planning A* **35**(10) 1785–1807

Dietzenbacher E, Günlük-Şenesen G, 2003b, "A structural decomposition analysis for a partially closed input-output model" paper presented at the 50th North American Meetings of the Regional Science Association International, Philadelphia; copy from E Dietzenbacher, University of Groningen, Groningen

Dietzenbacher E, Hoen A R, Los B, 2000, "Labor productivity in Western Europe 1975-1985: an intercountry, interindustry analysis" *Journal of Regional Science* **40**(3) 425–452

Dietzenbacher E, Los B, 1998, "Structural decomposition techniques: sense and sensitivity" *Economic Systems Research* **10**(4) 307–324

Dietzenbacher E, Los B, 2000, "Structural decomposition analysis with dependent determinants" *Economic Systems Research* **12**(4) 497–514

Dietzenbacher E, Los B, Stehrer R, Timmer M, de Vries G, 2013, "The construction of world input–output tables in the WIOD project" *Economic Systems Research* **25**(1) 71–98

Hermannsson K, Lisenkova K, Lecca P, Swales J K, McGregor P, 2014a, "The regional economic impact of more graduates in the labour market: a 'micro-to-macro' analysis for Scotland" *Environment and Planning A* **46**(2) 471–487

Hermannsson K, Lisenkova K, McGregor P, Swales J K, 2013, "The expenditure impacts of individual higher education institutions and their students on the Scottish economy under a regional government budget constraint: homogeneity or heterogeneity" *Environment and Planning A* **45**(3) 710–727

Hermannsson K, Lisenkova K, McGregor P, Swales J K, 2014b, "'Policy scepticism' and the impact of Scottish higher education institutions (HEIs) on their host region: accounting for the regional budget constraints under devolution" *Regional Studies* **48**(2) 400–417

Hewings G J D, Romanos M C, 1981, "Simulating less developed regional economies under conditions of limiting information" *Geographical Analysis* **13**(4) 373–390

Jensen-Butler C, Madsen B, 2005, "Decomposition analysis: an extended theoretical foundation and its application to the study of regional growth in Denmark" *Environment and Planning A* **37**(12) 2189–2208

Llop M, 2007, "Comparing multipliers in the social accounting matrix framework: the case of Catalonia" *Environment and Planning A* **39**(8) 2020–2029

McGregor P G, Swales J K, Yin Y P, 1999, "Spillover and feedback effects in general equilibrium interregional models of the national economy: a requiem for interregional input-output", in *Understanding and Interpreting Economic Structure* Eds. G J D Hewings, M Sonis, M Madden, Y Kimura (Springer, Berlin) pp 167–190

Miller R E, Blair P D, 2009 *Input-Output Analysis: Foundations and Extensions* 2nd edition (Cambridge University Press, Cambridge)

National Bureau of Statistics (NBS), 1999 *Input-Output Tables of China 1997* (China Statistics Press, Beijing)

National Bureau of Statistics (NBS), 2008 *China Statistical Year Book 2008* (China Statistics Press, Beijing)

National Bureau of Statistics (NBS), 2009 *Input-Output Tables of China 2007* (China Statistics Press, Beijing)

Oosterhaven J, Broersma L, 2007, "Sector structure and cluster economies: a decomposition of regional labour productivity" *Regional Studies* **41**(5) 639–659

Pei J, Dietzenbacher E, Oosterhaven J, Yang C, 2011, "Accounting for China's import growth: a structural decomposition for 1997-2005" *Environment and Planning A* **43**(12) 2971–2991

Pyatt G, 1988, "A SAM approach to modeling" *Journal of Policy Modeling* **10**(3) 327–352

Pyatt G, 2001, "Some early multiplier models of the relationship between income distribution and production structure" *Economic Systems Research* **13**(2) 139–163

Rose A, Casler S, 1996, "Input-output structural decomposition analysis: a critical appraisal" *Economic Systems Research* **8**(1) 33–62

Takayama A, 1985, *Mathematical Economics* 2nd edition (Cambridge University Press, Cambridge)

Trigg A B, Madden M, 1994, "Using a demand system to estimate extended input-output multipliers" *Economic Systems Research* **6**(4) 385–395

Schumann J, 1994, "Does it make sense to apply the static open input-output model for imputation and structural decomposition" *Economic Systems Research* **6**(2) 171–178

Yang C, Chen X, Xu J, 2008, "A method to optimize gross fixed capital investments for water conservancy in China" *Economic Systems Research* **20**(2) 151–172