A Generalized Hypothetical Extraction Analysis*

Umed Temurshoev†

January 29, 2009

Abstract

We explicitly formulate (optimization) problems of finding a *key sector* and a *key group* of sectors within the framework of a hypothetical extraction method (HEM), and derive their solutions in terms of simple measures termed industries' *factor worths*. It is shown that the top $k \geq 2$ sectors with the largest total contributions to some factor, in general, do *not* constitute the key group of k sectors, the issue which is totally ignored in the input-output linkage literature. The link to the fields of influence approach is discovered, which gives an alternative economic interpretation for the HEM problems in terms of sectors' input self-dependencies. Further, we examine how a change in an input coefficient affects the importance of an industry. The key group problem is applied to the Australian economy for factors of water use, CO_2 emissions, and generation of profits and wages.

Keywords: key group of sectors, hypothetical extraction, fields of influence, redundancy, input-output

JEL Classification Codes: C67, E61, L52, O21

*I am particularly grateful to Michael Alexeev, Erik Dietzenbacher and Bart Los for their constructive comments and suggestions. I also thank Geoffrey Hewings, Michael Lahr, Albert Steenge, Marcel Timmer and Shlomo Weber for their comments. I am grateful to Manfred Lenzen for providing necessary data on Australian economy. This paper also benefitted from presentations at the EERC Workshops (Kyiv, 2007, 2008), EUNIOUS Workshop (Groningen, 2008), and the 2-nd Sino-Netherlands Workshop on Input-Output Techniques (Beijing, 2008). I wish to thank the EERC for financial support.

†Faculty of Economics & Business and SOM Research School, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. Tel.: +31 50 363 8530; fax: +31 50 363 7337. E-mail: u.temurshoev@rug.nl

1 Introduction

There are ample studies within the input-output (IO) framework that investigate the issue of identification of so-called "key sectors" - sectors with the largest potential of spreading growth impulses throughout the economy. The issue of key sectors determination is seen to be useful for economic planning, in particular, in developing countries. From the development strategy point of view, it is reasonable for a country with a limited amount of financial resources to invest in those few industries, which have the largest impact on the whole economy through their buying and selling linkages with all other production units.[1] This approach, pioneered by Rasmussen (1956) and Hirschman (1958), was followed by a vast number of theoretical and empirical studies, and still constitutes one of the main areas in the IO and regional economics (see e.g., Strassert 1968, Yotopoulos and Nugent 1973, Jones 1976, Schultz 1977, Cella 1984, Hewings et al. 1989, Heimler 1991, Dietzenbacher 1992, Sonis et al. 1995, Dietzenbacher and van der Linden 1997, Cai and Leung 2004, Cardenete and Sancho 2006, Midmore et al. 2006, Beynon and Munday 2008).

However, the meaning of key sectors for economic development is rather debatable, since economic growth is determined not only by the structure and strength of inter-sectoral linkages, but also by production constraints, final demand and employment structure, imports, institutional and policy settings, income distribution, and technical and human capital endowment. Therefore, the application of key sector determination goes beyond examining only production linkages. For example, Diamond (1975), Meller and Marfán (1981), Groenewold et al. (1987, 1993) and Kol (1991) analyze employment linkages for Turkey, Chile, Australia, and for Indonesia, South Korea, Mexico and Pakistan, respectively. Gould and Kulshreshtha (1986) examine the impacts of final demand changes on energy use for Saskatchewan economy employing linkage analysis. Since according to the classical development economics for developing countries economic growth is intrinsically linked to changes in the structure of production, many studies applied the notion of key sectors for the analysis of structural change (see e.g., Hewings et al. 1989, Sonis et al. 1995,

[1]It is also true that the overall economic growth depends on the sectoral growth rates, which are in turn dependent on the linkages between the sectors. Strong linkages provide a possibility of gaining competitive advantage for industries. For instance, if a sector successfully enters a foreign market, it will be easier for industries (firms) that have high linkages with this sector to gain access to the foreign market as well (Porter 1990, Hoen 2002).

Roberts 1995). Given current concerns about the environmental problems, Lenzen (2003) focuses on economic structure of Australia in terms of resource use and pollutant emissions by identifying key sectors and linkages that have large environmental impacts in the form of resource depletion and ecosystem degradation.

In this paper we focus on the linkage analysis based on a hypothetical extraction method (HEM), which have become increasingly popular (Miller and Lahr 2001). Just to mention a few recent studies, the HEM has been applied to the analysis of water use (Duarte et al. 2004), key sectors identification (Andreosso-O'Callaghan and Yue 2004), the role of the agriculture sector (Cai and Leung 2004), the construction sector (Song et al. 2006) and the real estate sector (Song and Liu 2007). Los (2004) proposes to identify strategic industries using the HEM in a dynamic IO growth model. The HEM is also a useful tool to evaluate the significance of a sector in cases of crises-driven threats of industry shutdowns, which may help governments to decide whether to support financially the sector under threat or not.[2] The main contribution of this paper to the literature on key sectors identification from the HEM perspective is that we *explicitly* formulate the *optimization problems* of finding a *key sector* and a *key group of sectors*, and derive their solutions in terms of simple measures called industries' *factor worths*. The term "factor" refers to any indicator that interests an analyst in identifying the most important industries, which might be a social factor such as employment, income, government revenue, or an environmental factor such as primary energy consumption, greenhouse gas emissions, water use, land disturbance, or an economic/financial factor such as GDP, gross operating surplus, export/import propensity, or any combinations of these factors. The important implications of our formal formulation of the HEM are the following.

Firstly, given that we have found simple measures for quantifying industries' importance, an analyst does *not* have to perform a three-step procedure of the HEM (to be explained later), which becomes, in particular, a rather formidable task when the number of industries is rather large (say, 100 or more). Secondly, and more importantly, we distinguish between a *key sector problem* and a *key group problem* and show that the key group of $k \geq 2$ sectors is, in general, *different* from the set of

[2]The threat of downfall of the US car industry in the current financial crisis and debates on providing massive public spending to the industry can serve one such example. Other examples, are the downfall of the only Dutch aircraft manufacturer Fokker in 1995-96, and the disappearance of the Belgian national airline Sabena in 2001, both of which resulted in the shutdown of an entire national industry (Los 2004).

top k sectors selected on the base of the key sector problem. This is important, since up to date, to our best knowledge, the linkage literature accepted the top k sectors from the ranking of individual sector's contributions to economy-wide output as the key group. This incongruence is due to the fact that while the key sector problem looks for the effect of extraction of one sector, the key group problem considers the effect of a *simultaneous* extraction of $k \geq 2$ sectors that takes differently into account the cross-contributions of the extracted industries to total factor arising within and outside the group. This impact is largely dependent on the similarity/dissimilarity of the linkage pattern of sectors to each other. Thirdly, we show that the HEM is directly related to the fields of influence approach (Sonis and Hewings 1989, 1992), which gives an alternative economic interpretation of the HEM problems in terms of the overall impact on aggregate factor due to an incremental change in sectors' input self-dependencies. Finally, our formulation of the HEM allows to examine a *combined* key sector/group problem, where the objective is a combination of several factors. For instance, one may wish to identify a key sector that has simultaneously the largest total (direct and indirect) contribution to employment *and* the least total impact on carbon emissions generation.

We also examine the effect of a change in an input coefficient on the factor importance of an industry. It is shown that a positive (negative) change in a direct input coefficient a_{rc} never decreases (increases) the factor generating importance of any sector i, and surely increases (decreases) its factor worth if sector r requires directly and/or indirectly inputs from sector i. The economic interpretations of such change include, for example, an increase in complexity of technological links between sectors (or a rise in the density of the technology matrix), an increase in sectoral interdependence, innovation and technological progress, etc.

The rest of the paper is organized as follows. In Section 2.1 we present the optimization problem of finding a key sector, and examine how a change in a direct input coefficient affects the factor generating importance of industries. Section 2.2 generalizes the key sector problem to a key group identification problem, whose solution is defined in terms of a *group factor worth* of industries. The combined key sector/group problem is examined in Section 2.3. In Section 3 the link between the HEM and fields of influence method is explored. Section 4 contains results from the empirical application of the key sector and key group problems to the Australian economy. Section 5 concludes. All proofs are relegated to the Appendix.

2 Key sector problem vs. key group problem

2.1 Finding the key sector

The main point of departure is the open static Leontief model (see e.g., Miller and Blair 1985), given by $\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{f}$, where $\mathbf{x}$ is the $n \times 1$ endogenous vector of gross outputs of n sectors, $\mathbf{A}$ is the n-square direct input requirements matrix, and $\mathbf{f}$ is the $n \times 1$ exogenous vector of final demands (including consumption, investments, and government expenditures).[3] The input coefficients a_{ij} denote the output in industry i directly required as input for one unit of output in industry j, hence the ith element of the vector $\mathbf{A}\mathbf{x}$ gives the total *intermediate* demand of all sectors for the output of industry i. That is, the fundamental equation of the open Leontief system states that gross output, $\mathbf{x}$, is the sum of all intermediate demand, $\mathbf{A}\mathbf{x}$, and final demand, $\mathbf{f}$. The reduced form of the model is

$$\mathbf{x} = \mathbf{B}\mathbf{f}, \tag{1}$$

where $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$ is the Leontief inverse with $\mathbf{I}$ being the identity matrix. Its element b_{ij} denotes the output in industry i directly and indirectly required to satisfy one unit of final demand in industry j. The row vector of *output multipliers* is defined as $\mathbf{m}'_o = \boldsymbol{\imath}'\mathbf{B}$, where $\boldsymbol{\imath}$ is a summation vector consisting of ones. Its j-th element $m^o_j = \sum_{k=1}^{n} b_{kj}$ indicates the increase of total output in all industries per unit increase of final demand in industry j.

For the purpose of identification of important sectors we adopt the *hypothetical extraction method* (HEM) originally developed and used by Paelinck et al. (1965), Strassert (1968) (as cited in Miller and Lahr 2001) and Schultz (1977), the central idea of which is briefly as follows. To estimate the importance of sector i to the economy, delete the i-th row and column of the input matrix $\mathbf{A}$, and then using (1) compute the reduced outputs in this hypothetical case (the final demand vector also excludes f_i). The difference between *total* outputs of the economy before and after

[3]Adopting usual convention, matrices are given in bold, uppercase letters; vectors in bold, lowercase letters; and scalars in italic lowercase letters. Vectors are columns by definition, and transposition is indicated by a prime.

the extraction (called "total linkage") measures the relative stimulative importance of sector i to the economy.[4]

However, unlike the traditional HEM approach, we allow for a rather general definition of importance, which may be used to address various economic, social, or environmental issues.[5] For instance, key sectors may be determined according to their potential of generating income, emission of greenhouse gases, creating jobs, or resource use. For the purpose of a general exposition of the HEM problem, we refer to the various policy-relevant indicators as *factors*. Let the vector of *direct factor coefficients* $\boldsymbol{\pi}$ denotes the sectoral factor usage per unit of total output, hence the row vector of *factor multipliers* is $\mathbf{m}'_{\pi} = \boldsymbol{\pi}'\mathbf{B}$, and its j-th element $m^{\pi}_j = \sum_{k=1}^{n} \pi_k b_{kj}$ indicates the economy-wide increase of factor usage/production per unit increase of final demand in industry j.

We are now in a position to address the key sector identification problem. Let first denote by $\mathbf{A}^{-i}$ the new input matrix derived from $\mathbf{A}$ by setting to zero all of its i-th row and column elements. The crucial assumption made (which is usual for all the HEM approaches) is that in a new system without sector i the input structure of sectors $j \neq i$ remains unchanged. From economic point of view, this implies that foreign (external) industries substitute sector i in providing its output in order to satisfy the intermediate demand of the remaining industries and the final demand for commodity i. Although at first glance this assumption seems restrictive, in fact it is not given our main aim of identifying the importance of sector i. The point is that by taking all other input coefficients fixed, we explicitly allow the resulting outcome to depend only on extraction of sector i, which is now not participating in the "roundabout" of the production process. The vector of total outputs after

[4]This method was criticized for the reason that it does not distinguish the total linkages into backward and forward linkages (see e.g., Meller and Marfán 1981, Cella 1984, Clements 1990, Dietzenbacher and van der Linden 1997). However, we believe that for measuring a sector's economy-wide impact it is the most adequate HEM, since setting to zero only a column (row) to compute the backward (forward) linkages in the non-complete HEM takes only one-sided impact into account. Moreover, the last two linkage measures are closely related in the sense of the forward-link involvement problem of backward linkage measures, and, vice versa, the backward-link presence in the forward linkage measures (see e.g., Yotopoulos and Nugent 1973, Cai and Leung 2004). See Miller and Lahr (2001) for an excellent discussion on all possible extractions, who state that for the purpose of finding a key sector "we believe the original hypothetical extraction approach ... is totally adequate - Meller and Marfán and other modifications notwithstanding" (p. 429).

[5]For example, Ten Raa (2005, p. 26) states: "Output increases induced by a final demand stimulus are of little interest in themselves. What matters is the income generated by the additional economic activity."

extracting sector i is $\mathbf{x}^{-i} = \mathbf{B}^{-i}\mathbf{f}^{-i}$, where $\mathbf{B}^{-i} = (\mathbf{I} - \mathbf{A}^{-i})^{-1}$, and $\mathbf{f}^{-i}$ is the same as $\mathbf{f}$ except its i-th entry that is set to zero. The reason for excluding f_i in the final demand vector $\mathbf{f}^{-i}$ is that when sector i ceases to exist, its (domestic) output should be zero, which from (1) is equivalent to $f_i = 0$ (see also e.g., Schultz 1977, Miller and Lahr 2001).

The objective is picking the appropriate sector i, such that its extraction from the system generates the highest possible reduction in the factor of interest (say, total income). Formally, the problem is

$$\max\{\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-i} \,|i = 1, \ldots, n\}. \tag{2}$$

This is a finite optimization problem, which has at least one solution. A solution to (2) is denoted by i^* and is called the *key sector*. Removing i^* from the initial production structure has the largest overall impact on the factor generation. To solve (2) we use the following result due to Ballester et al. (2006, Lemma 1, p. 1411), our proof of which is given in the Appendix.[6]

Lemma 1. *Let* $\mathbf{B}$ *and* $\mathbf{B}^{-i}$ *be, respectively, the Leontief inverses before and after extraction of sector* i *from the production system, and* $\mathbf{e}_i$ *the* i*-th column of the identity matrix. Then* $\mathbf{B} - \mathbf{B}^{-i} = \frac{1}{b_{ii}}\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B} - \mathbf{e}_i\mathbf{e}_i'$.

Using Lemma 1 problem (2) after some mathematical transformations can be rewritten as (see Appendix):

$$\max\left\{\frac{1}{b_{ii}}\mathbf{m}_\pi'\mathbf{e}_i\mathbf{e}_i'\mathbf{x}\Big|i = 1, \ldots, n\right\}. \tag{3}$$

The problem in (2) is equivalent to $\min\{\boldsymbol{\pi}'\mathbf{x}^{-i} \,|i = 1, \ldots, n\}$. However, a direct use of one of these criteria in determining the key sector in empirical applications forces an analyst to extract different sectors separately and compute the required objective n times, which becomes a formidable task when n is large. Although with modern technology this is not a big issue, problem (3) shows that there exist a much

[6] We should note that Lemma 1 in Ballester et al. (2006) is given for a *symmetric* adjacency matrix in the social network framework, and does not consider the ii-th element of the difference $\mathbf{B}-\mathbf{B}^{-i}$. For asymmetric case, change $m_{ij}(\mathbf{g}, a)$ to $m_{ji}(\mathbf{g}, a)$ in their Lemma 1. Although Ballester et al. (2006) investigate identification of a key player in social networks, there is a direct link to the key sector problem (see for details Temurshoev 2009).

simpler way to get the desired outcome. We define the *factor worth* of sector i as

$$\omega_i^\pi(\mathbf{A},\mathbf{f},\boldsymbol{\pi}) = \frac{m_i^\pi x_i}{b_{ii}}.$$

Thus, given the objective in (3) we have established the following result.

Theorem 1. *The key sector i^* that solves $\max\{\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-i} \; | i = 1,\ldots,n\}$ has the highest factor worth, i.e., $\omega_{i^*}^\pi(\mathbf{A},\mathbf{f},\boldsymbol{\pi}) \geq \omega_i^\pi(\mathbf{A},\mathbf{f},\boldsymbol{\pi})$ for all $i = 1,\ldots,n$.*

From Theorem 1 it follows that the standard measure of the high factor multiplier m_i^π is not sufficient for sector i to be an optimal target, say, for investments. For the last, besides m_i^π, the size of the sector's output x_i and its self-dependency as indicated by b_{ii} are equally important, where the first has a positive effect, while the second an inverse effect on the worth of sector i.

The traditional gross output approach of the HEM corresponds to the problem (2) or (3) when a summation vector $\boldsymbol{\imath}$ is substituted for the vector of factor coefficients $\boldsymbol{\pi}$. The following result is then an immediate outcome of Theorem 1.

Corollary 1. *The key sector i^* that solves $\max\{\boldsymbol{\imath}'\mathbf{x} - \boldsymbol{\imath}'\mathbf{x}^{-i} \; | i = 1,\ldots,n\}$ has the highest output worth, i.e., $\omega_{i^*}^o(\mathbf{A},\mathbf{f}) \geq \omega_i^o(\mathbf{A},\mathbf{f})$ for all $i = 1,\ldots,n$, where $\omega_i^o(\mathbf{A},\mathbf{f}) = m_i^o x_i/b_{ii}$ is the (gross) output worth of sector i.*

Next we examine how larger interdependence of sectors affect the factor worth of sector i. Let the input matrix $\widetilde{\mathbf{A}}$ represent the more "complex" input structure than $\mathbf{A}$, and, without loss of generality, assume that $\widetilde{\mathbf{A}}$ differs from $\mathbf{A}$ only with respect to the rc-th element that is increased by $\alpha > 0$. Then it is apparent that $\widetilde{\mathbf{B}} = \mathbf{I} + \widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^2 + \cdots > \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots = \mathbf{B}$,[7] which in turn implies that, given $\mathbf{f}$ and $\boldsymbol{\pi}$, both the numerator and denominator in the definition of the factor worth of sector i might only increase, hence it is not clear whether $\omega_i^\pi(\widetilde{\mathbf{A}},\mathbf{f},\boldsymbol{\pi})$ is larger or smaller than $\omega_i^\pi(\mathbf{A},\mathbf{f},\boldsymbol{\pi})$. Nevertheless, in Theorem 2 below we are able to show that a rise in direct input interdependence between two sectors never decreases sector i's factor worth, and, moreover, we establish a necessary and sufficient condition under which such a change surely increases $\omega_i^\pi(\widetilde{\mathbf{A}},\mathbf{f},\boldsymbol{\pi})$.

Theorem 2. *Let the input matrix $\widetilde{\mathbf{A}}$ differs from $\mathbf{A}$ only with respect to the rc-th entry, which has changed by $\alpha \neq 0$. Given $\mathbf{f}$ and $\boldsymbol{\pi}$, if $\alpha > 0$ (resp. $\alpha < 0$) then*

[7] We write $\mathbf{X} > \mathbf{Y}$ if $x_{ij} \geq y_{ij}$ for all i, j, with at least one strict inequality.

$\omega_i^{\pi}(\widetilde{\mathbf{A}}, \mathbf{f}, \boldsymbol{\pi}) \geq \omega_i^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi})$ (resp. $\omega_i^{\pi}(\widetilde{\mathbf{A}}, \mathbf{f}, \boldsymbol{\pi}) \leq \omega_i^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi})$) for all $i = 1, \ldots, n$, with equality holding if and only if $b_{ir} = 0$.

One implication of Theorem 2 is that when the technology becomes more complex in a sense that domestic industries become more interdependent on each other, then the factor generating importance of any sector never falls. Moreover, sector i's worth surely increases if $b_{ir} > 0$, i.e., when the supplying sector r, whose product were demanded more per unit of output of sector c, has direct and/or indirect input requirements from industry i. Second implication is that more efficient technology never increases the factor worth of any sector for the same vectors of final demand and factor coefficients. In particular, if, say, due to innovation a_{rc} decreases, then sector i's importance weakens whenever $b_{ir} > 0$. This is because now sector c is less dependent on sector r, which in its turn uses inputs (directly and/or indirectly) from sector i, thus the first weaker dependence affects the worth of sector i negatively.

The straightforward special case of Theorem 2 is when $\boldsymbol{\pi} = \boldsymbol{\imath}$, which shows that the output worth of sector i increases (decreases) if the input coefficient a_{rc} increases (decreases) and sector r purchases inputs directly and/or indirectly from industry i.

Corollary 2. *Assume that the input coefficient a_{rc} changes by $\alpha \neq 0$, i.e., $\tilde{a}_{rc} = a_{rc} + \alpha$. Given $\mathbf{f}$, if $\alpha > 0$ (resp. $\alpha < 0$), then $\omega_i^{o}(\widetilde{\mathbf{A}}, \mathbf{f}) \geq \omega_i^{o}(\mathbf{A}, \mathbf{f})$ (resp. $\omega_i^{o}(\widetilde{\mathbf{A}}, \mathbf{f}) \leq \omega_i^{o}(\mathbf{A}, \mathbf{f})$) for all $i = 1, \ldots, n$, with equality holding if and only if $b_{ir} = 0$.*

2.2 Finding the key group

Although the linkage literature using the HEM acknowledges the possibility of extraction of several industries, the theoretical analysis does not go beyond describing it using partitioned matrices to the reduced form of the Leontief model (see e.g., Miller and Lahr 2001). This, however, is quite complex to implement empirically since one has to consider all possible combinations of certain number of industries from totality of n sectors (correspondingly changing the members and nonmembers of partitioned matrices) in order to determine the most important group of sectors, which explains why there is no any empirical study that explicitly focuses on the role of several industries simultaneously. Hence, in all studies, to our best knowledge, the HEM was applied to only one sector, and the most important industries were defined to be those with the largest individual contributions to total output (or any other factor).

In this section we wish to fill this gap in the literature, generalizing the key sector problem from the previous section to the *key group problem.* Similar to the notion of individual key sector, a *key group* of $k \geq 2$ sectors is defined as the group of industries, whose removal from the production system has the largest impact on the factor consumption/generation.[8] Since the two problems are inherently different, we expect that, in general, the top k sectors with the largest factor worths do *not* compose the *key group*, which is also confirmed in the empirical application in Section 4. The underlying reason for this outcome is that industries can be *redundant* (or, equivalently, similar to each other) with respect to their linkage patterns to other sectors and their capabilities of factor generation. Hence, targeting industries with the same linkage characteristics might not be an optimal policy strategy, but instead choosing sectors with *heterogenous* linkage structures will have the largest impact on the factor usage/generation.

The objective is now picking k $(1 \leq k \leq n)$ sectors $i_1, i_2, \ldots, i_k$ $(i_s \neq i_r)$ such that their extraction from the production structure generates the largest impact on the overall factor consumption/generation, i.e.,

$$\max\{\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-\{i_1,\ldots,i_k\}} \,|i_1, \ldots, i_k = 1, \ldots, n; i_s \neq i_r\}, \tag{4}$$

where $\mathbf{x}^{-\{i_1,\ldots,i_k\}} = \mathbf{B}^{-\{i_1,\ldots,i_k\}}\mathbf{f}^{-\{i_1,\ldots,i_k\}}$, and the superscript $-\{i_1, \ldots, i_k\}$ refers to the situation when sectors $i_1, i_2, \ldots, i_k$ are hypothetically extracted from the economy. Note that along the similar reasonings made in Section 2.1, $\mathbf{f}^{-\{i_1,\ldots,i_k\}}$ is exactly the same as $\mathbf{f}$ but with $f_{i_s} = 0$ for all $s = 1, \ldots, k$. The solution to (4) is denoted by $\{i_1^*, i_2^*, \ldots, i_k^*\}$ and is called the *key group of size k*.

The following important identity characterizes the changes in all elements of the Leontief inverse when a group of k sectors is hypothetically extracted from the production system.

Lemma 2. *Let $\mathbf{B}^{-\{i_1,\ldots,i_k\}}$ be the Leontief inverse after extraction of sectors $i_1, i_2, \ldots, i_k$ from the production system, where $1 \leq k \leq n$. Then the identity $\mathbf{B} - \mathbf{B}^{-\{i_1,\ldots,i_k\}} =$*

[8] Note that if the factor generation is unfavorable from societal point of view (e.g., an increase in CO_2 emissions has detrimental consequences) and the policy-makers want to find the *least* harmful industries to target on, then the key group will be defined as those industries that have the *smallest* impact of the factor generation.

$\mathbf{BE}\left(\mathbf{E}'\mathbf{BE}\right)^{-1}\mathbf{E}'\mathbf{B} - \mathbf{EE}'$ *always holds, where* $\mathbf{E}$ *is the* $n \times k$ *matrix defined as* $\mathbf{E} = (\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_k})$.

Note that Lemma 1 is just a special case of Lemma 2 with $k = 1$. We should also note that the k extracted sectors can be arbitrary ordered, hence the matrix $\mathbf{E}$ can have different ordering of the identity columns corresponding to those sectors.[9] Using Lemma 2 it can be shown that the problem (4) is exactly equivalent to (see Appendix)

$$\max\left\{\mathbf{m}_\pi'\mathbf{E}(\mathbf{E}'\mathbf{BE})^{-1}\mathbf{E}'\mathbf{x}\,|i_1, \ldots, i_k = 1, \ldots, n; i_s \neq i_r\right\}. \tag{5}$$

Note that in the maximization process the vectors of factor multipliers and gross outputs and the Leontief inverse matrix (i.e., $\mathbf{m}_\pi$, $\mathbf{x}$ and $\mathbf{B}$) are all given, and only the k identity columns in $\mathbf{E}$ are changed in order to consider all possible combinations of k sectors from all n industries. Now define the *group factor worth* of sectors $i_1, \ldots, i_k$ ($i_r \neq i_s$) as

$$\omega^\pi_{i_1,\ldots,i_k}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) = \mathbf{m}_\pi'\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{x},$$

where $\mathbf{B}_{kk} = \mathbf{E}'\mathbf{BE}$, which includes all the elements of the Leontief inverse that are directly related to the extracted sectors.

Given the key group problem (5), we thus have the following result.

Theorem 3. *For* $1 \leq k \leq n$ *the key group of size* k $\{i_1^*, i_2^*, \ldots, i_k^*\}$ *that solves* $\max\{\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-\{i_1,\ldots,i_k\}}\,|i_1, \ldots, i_k = 1, \ldots, n; i_s \neq i_r\}$ *has the highest group factor worth, i.e.,* $\omega^\pi_{i_1^*,\ldots,i_k^*}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) \geq \omega^\pi_{i_1,\ldots,i_k}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi})$ *for all* $i_1, \ldots, i_k = 1, \ldots, n$ *with* $i_s \neq i_r$.

Note that the key group problem in (5) with $k = 1$ boils down to the key sector problem (3), hence given the definition of the group factor worth, Theorem 1 is also a particular case of Theorem 2 when the target is only one sector (i.e., $k = 1$).

When the key group of size k is searched in the spirit of the traditional HEM approach, Theorem 3 implies the following result.

[9] Notice that if $k = n$ and $\mathbf{E} = \mathbf{I}$, then $\mathbf{B} - \mathbf{B}^{-\{i_1,\ldots,i_k\}} = \mathbf{B} - \mathbf{I}$, which is expected. However, in this case also $\mathbf{E}$ does not have to be an identity matrix, but $\mathbf{E}$ being any permutation matrix of order n gives the desired result.

Corollary 3. *For* $1 \leq k \leq n$ *the key group of size* k $\{i_1^*, \ldots, i_k^*\}$ *that solves* $\max\{\boldsymbol{\imath}'\mathbf{x} - \boldsymbol{\imath}'\mathbf{x}^{-\{i_1,\ldots,i_k\}} \mid i_1, \ldots, i_k = 1, \ldots, n; i_s \neq i_r\}$ *has the highest group output worth, i.e.,* $\omega^o_{i_1^*,\ldots,i_k^*}(\mathbf{A}, \mathbf{f}) \geq \omega^o_{i_1,\ldots,i_k}(\mathbf{A}, \mathbf{f})$ *for all* $i_1, \ldots, i_k = 1, \ldots, n$ *with* $i_s \neq i_r$ *and* $\omega^o_{i_1,\ldots,i_k}(\mathbf{A}, \mathbf{f}) = \mathbf{m}_o'\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{x}$.

While the key sector problem looks for the effect of extraction of one sector, the key group problem considers the effect of a *simultaneous* extraction of $k \geq 2$ sectors. This implies that the two problems are not equivalent since the key group problem takes into full account all the cross-contributions of the extracted sectors to the overall factor that is used/generated both within and outside the group. For example, if two industries are perfectly identical with respect to their linkages patterns (including input coefficients' sizes) and more or less also similar in terms of their final demand and factor generation structure, then their group worth is expected to be less than that of the group, which consists of one of the mentioned sectors together with another industry that has quite different patterns of (significant) interindustry linkages and factor generation ability. The *redundancy principle* is well-known in the sociology literature on social networks that emphasizes the redundancy of actors with respect to adjacency, distance, and bridging (see e.g., Burt 1992, Borgatti 2006). Arguing that the information and control benefits of a large and *diverse* network are more than those of a small and homogeneous network, Burt (1992, p.17), for example, states: "What matters is the number of nonredundant contacts. Contacts are redundant to the extent that they lead to the same people, and so provide the same information benefits." Taking redundancy into account is crucial in determining the most important group in social networks (see Everett and Borgatti 1999, 2005, Temurshoev 2008). In general, within the IO framework, we expect that k (≥ 2) sectors with the largest individual factor worths will not be much different from the key group of size k only if the IO tables are highly aggregated. Otherwise, the difference should be in place, and will largely depend on the structures of the production system, direct factor coefficients and gross outputs.

2.3 The combined key sector/group problem

Nowadays policy-makers, governments, companies and the general public are all becoming engaged with the phenomenon of "sustainability", which was brought to the public attention by environmental movements about 30 years ago that mainly emphasizes the issue of some sort of tradeoff between economic development and environmental quality. Hence, the concept of sustainable development is becoming the main focus, which requires meeting increasing environmental concerns along with maintaining economic development. For this reason corporations are beginning to be more and more involved in using the so-called *triple bottom line* (TBL) accounting through which economic, social and environmental spheres of sustainability are assessed and reported (see e.g., Henriques and Richardson 2004). Further, at the country level Foran et al. (2005) develop a numerate TBL account of the Australian economy with ten indicators that accounts for the full supply chain approach using the generalized IO analysis, against which many management issues at lower (say, firms) levels can be benchmarked.

The generalized HEM approach proposed in this paper can be applied to the sustainable development policy design and analysis.[10] The key group problem (4) can easily accommodate the notion of TBL approach from the HEM perspective. Let take the economic, social and environmental factors in the example of value-added, employment and CO_2 emissions, respectively. If $\mathbf{v}$, $\mathbf{l}$ and $\mathbf{c}$ denote, respectively, the direct value-added, labor and CO_2 coefficients vectors, the total (direct and indirect) value-added, employment and CO_2 emissions that is generated to satisfy the final demand $\mathbf{f}$ is equal to $\mathbf{v}'\mathbf{x}$, $\mathbf{l}'\mathbf{x}$ and $\mathbf{c}'\mathbf{x}$, correspondingly. Then a *combined key sector* and a *combined key group* problems are given, respectively, by problems (2) and (4) with the direct factor coefficients defined as $\boldsymbol{\pi} = \mathbf{v} + \mathbf{l} - \mathbf{c}$. Note that since CO_2 generation is unfavorable, its direct coefficients are entered with a minus sign in the definition of $\boldsymbol{\pi}$. Also notice that factors written is this form can have an economic meaning only if they are all expressed in the same measurement unit. This can be done, for example, by multiplying the number of jobs by a price so that employment is expressed in some currency term (like in the index number literature). Or, one

[10] An example of such policy is given by Daniels (1992): since the 1980s Australia has expanded its exports of meat, wool, wheat and non-ferrous metals to maintain revenues and living standards in response to increasing foreign debts and falling primary commodity prices. However, since these exports are highly environmental damaging activities, "Australia became locked into an environmental-economic dilemma through increasing dependency on degrading production and further erosion of environmental quality. Daniels argued that, in order to avoid long-term losses of productivity, biodiversity and real income, Australia has to re-direct its domestic production towards more value-adding and less land- and emissions-intensive commodities" (Lenzen 2003, p. 29).

might assign appropriate weights to each factor that is included in the combined factor coefficients vector. For instance, we may write $\mathbf{l} = t_v \mathbf{j}$, where the (number of) jobs coefficients $\mathbf{j}$ is expressed in terms of currency using the weight $t_v = \frac{\mathbf{v}'\mathbf{x}}{\mathbf{j}'\mathbf{x}}$ that indicates the value of value-added per one (full-time) job. Theorems 1 and 3 are then similarly used to identify the key sector and the key group of certain size in these combined problems.

3 The link to the fields of influence approach

Another well-known technique for evaluating sectors' influence on the rest of the economy is Sonis and Hewings' notion of a *field of influence* method (see e.g., Sonis and Hewings 1989, 1992). This methodology answers the question of how changes in some elements of the input matrix affect the rest of the system by examining the impact on the elements of the Leontief inverse, and is general enough to handle changes in one direct coefficient, in all elements of a row or column of the input matrix, or in all coefficients simultaneously.[11] To briefly introduce this method, let consider a change of $\alpha \neq 0$ in only one coefficient a_{rc}, with all other input coefficients being fixed. Then the Leontief inverse after the change is[12]

$$\widetilde{\mathbf{B}} = \mathbf{B} + \frac{\alpha}{1 - \alpha b_{cr}} \mathbf{F}(r, c), \tag{6}$$

where $\mathbf{F}(r, c) = \mathbf{B}\mathbf{e}_r \mathbf{e}_c' \mathbf{B}$ is the *first-order field of influence* matrix of the coefficient a_{rc} and $\mathbf{e}_r$ is the r-th column of the identity matrix. The sum of all elements of the first-order field of influence matrix, $\boldsymbol{\imath}'\mathbf{F}(r, c)\boldsymbol{\imath}$, gives the *first-order intensity field of influence* of the direct input a_{rc}. In Sonis and Hewings (1989) this concept was introduced in order to measure the *inverse importance* of direct inputs. Consequently, those elements of $\mathbf{A}$ whose changes lead to the largest impact on the system are called the *inverse-important coefficients*.

Unlike the standard first-order intensity $\boldsymbol{\imath}'\mathbf{F}(r, c)\boldsymbol{\imath}$, the scalar $\boldsymbol{\imath}'\mathbf{F}(r, c)\mathbf{f}$ weights every purchasing sector in the sum according to the size of its final demand, hence

[11] From economic point of view this enables one to analyze the effect of technological change, improvements in efficiency, changes in product lines, changes in the structure and complexity of an economy over time, changes in trade dependency of a country, etc.

[12] Notice that $\frac{\partial \tilde{b}_{ij}}{\partial \alpha}\big|_{\alpha=0} = f_{ij}(r, c) = b_{ir} b_{cj} = f_{cr}(j, i)$. Also the coordinate form of (3) is the well-know Sherman and Morrison (1950) formula of inverse change as $\tilde{b}_{ij} = b_{ij} + \alpha b_{ir} b_{cj} / (1 - \alpha b_{cr})$.

can be called the *output first-order intensity weighted field of influence* of a_{rc}. This makes more sense in computing the global intensity since every sector is not given an equal importance, but rather its scale of final demand satisfaction is taken into account. More generally, we term the scalar $\boldsymbol{\pi}'\mathbf{F}(r,c)\mathbf{f}$ as a *factor first-order intensity weighted field of influence* of the coefficient a_{rc}, since the last measures the effect of an input change on total factor generation rather than gross output. Having defined this intensity measure, we can rewrite the factor worth of sector i as

$$\omega_i^{\pi}(\mathbf{A},\mathbf{f},\boldsymbol{\pi}) = \frac{m_i^{\pi} x_i}{b_{ii}} = \frac{\boldsymbol{\pi}'\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B}\mathbf{f}}{b_{ii}} = \frac{\boldsymbol{\pi}'\mathbf{F}(i,i)\mathbf{f}}{b_{ii}},$$

which clearly shows that the key sector problem (2) searches for such sector i that, on the one hand, has a large economy-wide impact on total factor usage/generation due to (incremental) change in its *direct input self-dependency*, and on the other hand, is less input dependent on itself directly and indirectly. The first statement is true since the effect of a change in direct input self-dependency of sector i on the overall factor consumption/generation is given by the factor first-order intensity weighted field of influence of input coefficient a_{ii}, $\boldsymbol{\pi}'\mathbf{F}(i,i)\mathbf{f}$.

Next, using the fact that for a nonsingular matrix $\mathbf{X}$ the identity $\begin{vmatrix} \mathbf{X} & \mathbf{b} \\ \mathbf{c}' & 0 \end{vmatrix} = -|\mathbf{X}|(\mathbf{c}'\mathbf{X}^{-1}\mathbf{b})$ holds, where $|\mathbf{X}|$ is the determinant of $\mathbf{X}$, we can write the hl-th element of the matrix $\mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{B}$ (recall that $\mathbf{B}_{kk} = \mathbf{E}'\mathbf{B}\mathbf{E}$) as

$$\mathbf{b}_{h\bullet}'\mathbf{B}_{kk}^{-1}\mathbf{b}_{\bullet l} = \frac{-\begin{vmatrix} \mathbf{B}_{kk} & \mathbf{b}_{\bullet l} \\ \mathbf{b}_{h\bullet}' & 0 \end{vmatrix}}{|\mathbf{B}_{kk}|},$$

where $\mathbf{b}_{h\bullet}'$ is the h-th row of the matrix $\mathbf{B}\mathbf{E}$ and $\mathbf{b}_{\bullet l}$ is the l-th column of $\mathbf{E}'\mathbf{B}$. The numerator in the last equation is nothing else as the hl-th element of the *matrix field of influence of order k* of the input coefficients $a_{i_1 i_1}, a_{i_2 i_2}, \ldots, a_{i_k i_k}$, $\mathbf{F}[(i_1,i_1),(i_2,i_2),\ldots,(i_k,i_k)]$ (see e.g., Fritz et al. 2002).[13] Hence, the group fac-

[13]We should note that the only difference comes in signs when k is even, i.e., in the fields of influence approach the determinant in the numerator of the last equation is multiplied by $(-1)^k$. However, we believe that in our setting it should be always multiplied by minus, otherwise the elements will be negative, which then contradict the Leontief inverse property.

tor worth of sectors $i_1, \ldots, i_k$ $(i_r \neq i_s)$ can be rewritten as

$$\omega^{\pi}_{i_1,\ldots,i_k}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) = \mathbf{m}'_{\pi}\mathbf{E}\mathbf{B}^{-1}_{kk}\mathbf{E}'\mathbf{x} = \boldsymbol{\pi}\mathbf{B}\mathbf{E}\mathbf{B}^{-1}_{kk}\mathbf{E}'\mathbf{B}\mathbf{f} = \frac{\boldsymbol{\pi}'\mathbf{F}[(i_1, i_1), \ldots, (i_k, i_k)]\mathbf{f}}{|\mathbf{B}_{kk}|}.$$

This implies that the key group problem (4) searches for a group of k sectors with the highest group factor worth, which is directly proportional to the impact on overall factor generation of an incremental changes in *direct input self-dependencies* of sectors comprising the group,[14] and inversely related to their unit *own net input dependence* that excludes the indirect role of the group members. To see the interpretation of the second effect, let consider the group of size two. Then $|\mathbf{B}_{22}| = b_{ii}b_{jj} - b_{ij}b_{ji}$, which gives the *net* input dependence per unit of output of sectors i and j $(\neq i)$ on themselves. This follows since $b_{ij}b_{ji}$ (or, equivalently, $b_{ji}b_{ij}$) gives the total input requirements of sector i (j) on itself through sector j (i), and excluding this from the total own dependence of sectors i and j, $b_{ii}b_{jj}$, gives the unit own input dependence through other sectors $k \neq i, j$.[15]

All in all, we have shown that the (generalized) HEM and the fields of influence approach are closely related, which is not surprising since both methods deal with the same issue of the impact of a change in input coefficients on the entire economic system.

4 Application to the Australian economy

We have already noted that the input-output linkage studies (implicitly) accepted the k sectors (where $1 < k < n$) with the largest individual factor worths as the *key group* of k sectors. In this section by the example of the Australian economy we show that this is not true as long as the HEM approach is concerned, i.e., the k sectors with the highest factor worths, in general, do *not* compose the *key group* of size k.

We have used data from Foran et al. (2005) and Centre for Integrated Sustain-

[14]This interpretation is due to the economic meaning of $\boldsymbol{\pi}'\mathbf{F}[(i_1, i_1), \ldots, (i_k, i_k)]\mathbf{f}$, which we might similarly term as a *factor intensity weighted field of influence of order* k of input coefficients $a_{i_1 i_1}, \ldots, a_{i_k i_k}$.

[15]In case of three sectors, one may write $|\mathbf{B}_{33}| = b_{kk}(b_{ii}b_{jj} - b_{ij}b_{ji}) - b_{jk}(b_{ii}b_{kj} - b_{ki}b_{ij}) - b_{ik}(b_{jj}b_{ki} - b_{kj}b_{ji})$ for all $i \neq j \neq k \neq i$ which has the same interpretation of the net own input dependence of sector k. Other orderings of rows (and columns) of $\mathbf{B}_{kk}$ give similar interpretation for the other two sectors i and j.

ability Analysis (2005) that include the 1994-1995 Australian IO tables and satellite accounts at 136 industry-level classification.[16] For simplicity, the industries were codified, whose list is given in Table 2. The key sector/group problem is performed for two environmental, one financial and one social factors, which are, respectively, water use, carbon dioxide (CO_2) emissions, gross operating surplus, and wages and salaries. The results are reported in the first five columns of Table 1 in terms of *relative* group worths, i.e., the group factor worths as a percentage of the overall factor use/generation before the extraction of sectors comprising the group.[17] Hence, these relative measures refer to the percentage decrease in economy-wide factor use/generation caused by the extraction. We only report the top 5 groups of size $k \in [1, 4]$, and, obviously, the group with rank 1 in each list is the corresponding *key group*.

Several observations can be made from Table 1. The first and most obvious observation is that different objectives give different composition of the key group of certain size and different rankings of sectors or group of sectors. This is totally expectable, as different sectors perform different functions in the economy, thus should not be equivalent it terms of various factors consumption/production.

Second outcome is that the composition of the key group of size k is, in general, different from the k sectors with the largest (individual) factor worths, which confirms our expectation that the key sector problem is not equivalent to the key group problem. For example, let us look at the key group problem in terms of water use. The first column of Table 1 shows that Dairy cattle & milk (Dc) is the key sector in water use with the relative water consumption worth of 19.5%.[18] The key group of size two consists of the key sector Dc and Beef cattle (Bc) jointly accounting for 37.6% of the economy-wide water consumption, which, however, does *not* include Diary products (Dp) that has the second largest water (usage) worth. Further, the key group of size 3 besides Dc and Bc includes Water supply, sewerage and drainage services (Wa), which has only the sixth rank according to the key sector problem with water worth of 10.6% (not shown in Table 1). The traditional "top-list" ap-

[16]Foran et al. (2005) give detail description of the data sources and its construction.

[17]For instance, the relative profits (gross operating surplus) worth of sectors i and j ($\neq i$) equals $(\omega^p_{i,j}(\mathbf{A}, \mathbf{f}, \mathbf{p})/\mathbf{p}'\mathbf{x}) \times 100$, where $\mathbf{p}$ is the vector of sectoral direct profits coefficients, thus $\mathbf{p}'\mathbf{x}$ is the total gross operating surplus in the economy.

[18]In the language of the HEM problem, if Dairy cattle & milk (Dc) sector would be eliminated from the economy then the overall use of water would be reduced by 19.5%.

Table 1: Relative group factor worths of Australian industries, 1994-1995

Rank	Group of size k and its relative factor worth (%) $k=1$	$k=2$	$k=3$	$k=4$	Factor multipliers	Factor use/ generation	Factor responsibility
	Objective: Water use				(ths.l/A$)	(Tl)	(Tl)
1	Dc (19.5)	Bc, Dc (37.6)	Bc, Dc, Wa (48.1)	Bc, Dc, Vf, Wa (58.0)	Ri (7.47)	Dc (3.54)	Dp (2.89)
2	Dp (18.6)	Dc, Mp (37.3)	Dc, Mp, Wa (47.7)	Dc, Mp, Vf, Wa (57.5)	Sc (1.64)	Bc (3.23)	Mp (2.68)
3	Bc (18.2)	Bc, Dp (36.8)	Bc, Dc, Vf (47.6)	Bc, Dp, Vf, Wa (57.1)	Dc (1.48)	Wa (2.02)	Fd (1.35)
4	Mp (18.1)	Dp, Mp (36.4)	Bc, Dp, Wa (47.3)	Dp, Mp, Vf, Wa (56.7)	Su (1.26)	Vf (1.80)	Ho (1.13)
5	Vf (10.7)	Dc, Wa (30.0)	Dc, Mp, Vf (47.2)	Bc, Dc, Fd, Wa (55.9)	Bc (0.73)	Ri (1.43)	Wa (1.12)
	Objective: CO_2 emissions				(kg/A$)	(Mtonnes)	(Mtonnes)
1	El (32.8)	Bc, El (52.9)	Bc, Fr, El (64.4)	Bc, Fr, El, Is (69.0)	Fr (98.3)	El (136.6)	Mp (59.7)
2	Bc (20.3)	El, Mp (50.6)	Fr, Mp, El (62.1)	Bc, Fr, El, Wt (67.7)	Sw (25.2)	Bc (81.2)	El (53.8)
3	Mp (18.3)	El, Fr (44.9)	Bc, Is, El (57.6)	Bc, Fr, El, Rb (67.5)	Bc (17.9)	Fr (50.9)	Fr (38.0)
4	Fr (12.3)	El, Is (37.5)	Bc, El, Wt (56.4)	Bc, Fr, El, At (67.2)	Hw (15.4)	Is (17.9)	Rt (21.7)
5	Is (5.5)	El, Wt (36.4)	Bc, El, Rb (56.3)	Bc, Fd, Fr, El (67.1)	Lm (14.8)	At (10.1)	Rb (16.8)
	Objective: Gross operating surplus (profits)				(A$/A$)	(A$ Bln)	(A$ Bln)
1	Dw (21.7)	Dw, Wt (31.2)	Dw, Rb, Wt (37.1)	Dw, Rb, Rt, Wt (42.5)	Dw (0.84)	Dw (38.7)	Dw (41.6)
2	Wt (9.7)	Dw, Rb (27.9)	Dw, Rt, Wt (36.7)	Dw, Nb, Rb, Wt (41.7)	Si (0.68)	Wt (7.5)	Rb (11.9)
3	Rb (6.6)	Dw, Rt (27.5)	Dw, Nb, Wt (35.9)	Dw, Nb, Rt, Wt (41.3)	Bl (0.63)	Rb (7.1)	Rt (11.0)
4	Rt (5.9)	Dw, Ms (26.9)	Dw, Ms, Wt (35.1)	Dw, Ho, Rb, Wt (40.8)	Br (0.622)	St (6.44)	Wt (9.4)
5	Ms (5.3)	Dw, Nb (26.7)	Dw, Ho, Wt (35.0)	Dw, Ms, Rb, Wt (40.8)	Ng (0.62)	Ms (6.39)	Nb (9.0)
	Objective: Net wages and salaries				(A$/A$)	(A$ Bln)	(A$ Bln)
1	Wt (12.4)	Rt, Wt (22.8)	Hs, Rt, Wt (31.7)	Ed, Hs, Rt, Wt (40.5)	Ed (0.61)	Ed (14.6)	Rt (18.3)
2	Rt (10.9)	Hs, Wt (21.3)	Ed, Rt, Wt (31.6)	Gv, Hs, Rt, Wt (39.0)	Gd (0.58)	Hs (14.2)	Hs (15.5)
3	Hs (9.1)	Ed, Wt (21.24)	Ed, Hs, Wt (30.2)	Gv, Ed, Rt, Wt (38.8)	Hs (0.533)	Rt (11.7)	Ed (14.5)
4	Ed (9.09)	Hs, Rt (20.1)	Gv, Rt, Wt (30.1)	Hs, Nb, Rt, Wt (38.0)	Os (0.53)	Wt (11.6)	Gv (11.5)
5	Gv (7.8)	Ed, Rt (20.0)	Ed, Hs, Rt (29.1)	Ed, Nb, Rt, Wt (37.9)	Gv (0.50)	Gv (10.0)	Nb (10.8)
Tot.	136	9,180	410,040	13,633,830	136	136	136

Note: "Tot." is the total number of all possible groups of size k. Mathematically, it is equal to the combinations of $n = 136$ sectors taken k at a time, $C_k^n = n!/(k!(n-k)!)$. One teraliter (Tl) is equivalent to 10^{12} litres. The source of the seventh column "Factor use/generation" is the satellite accounts in Foran et al. (2005) and Centre for Integrated Sustainability Analysis (2005), while the rest are own computations based on these data. One megatonnes (Mtonnes) equals 10^6 tonnes. Sectors' abbreviations are listed in Table 2.

proach would consider the "key" group of size 4 consisting of dairy and beef cattle, and dairy and meet products (i.e., Dc, Dp, Bc and Mp as the top 4 sectors), while the formal key group problem finds beef and diary cattle (Bc, Dc), Vegetable and fruit growing (Vf), and Water supply, sewerage & drainage (Wa) to be the part of the key group. The legitimate question is why the "top-list" approach does not give the true outcome identified by the key group problem.[19] The group factor worth of sectors $i_1, \ldots, i_k$ can be rewritten as

$$\omega^{\pi}_{i_1,\ldots,i_k}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) = \sum_{s=i_1}^{i_k} \pi_s x_s + \sum_{j \neq i_1,\ldots,i_k} \pi_j \left(x_j - x_j^{-\{i_1,\ldots,i_k\}} \right),$$

[19]Note that in our example these two approaches give identical results for $k \in [1,4]$ when the objectives are profits, and wages and salaries. We should, however, stress that these observations by no means can subside the existence of the difference between the two approaches, and thus the key group problem should always be given preference over the "top-list" approach whenever the HEM is the study methodology. Application to the Kyrgyzstan economy for value-added and gross output resulted in a dramatic difference between the "top-list" and key group problem approaches in defining the key group (these results are not shown here as we have decided to focus only on the Australian economy).

which shows that the factor worth of the extracted sectors includes not only their *direct* contributions to factor usage/generation (the first sum), but also their *indirect* contributions to factor consumed/generated by every other sector outside the group (the second sum).[20] Hence, with inherently different structure and sizes of inter-sectoral links, intermediate and final demands, the group of k sectors will play quite a different role in overall factor usage/generation process than a single industry, in particular, through its indirect channel.

This result wedges a bridge between the IO linkage analysis and the sociology literature on actors' importance in social networks. This link has to do with what sociologists call a *redundancy principle* (see e.g., Burt 1992, Borgatti 2006), which in our framework means that sectors may be redundant with respect to their linkage patterns. That is, sectors can be redundant when they connect the same third industries to each other, or when they are connected to the same third parties, in both cases with approximately the same sizes of inter-industry transactions and gross outputs. Sectors are called to be *structurally equivalent* in the latter case of redundancy in the sociological terminology. In the framework of social networks, Temurshoev (2008) extended a notion of *intercentrality measure* introduced by Ballester et al. (2006) in identifying a *key player* from social planner perspective to a *group intercentrality measure*, and showed that there is a link between the key group members and clusters of similar agents, where clusters are identified by a hierarchical agglomerative cluster analysis. That is, the key group of actors contains members from different clusters, i.e., key group members are rather nonredundant with respect to the patterns of ties to their alters. We believe that namely this redundancy principle in the IO framework explains the fact that Dairy products (Dp) that ranks high in the key sector problem (i.e., for $k = 1$) is not contained in key groups of size $k > 1$ in Table 1 in case of water usage. For example, key group of size 2 contains dairy and beef cattle (Dc and Bc) and not the second largest consumer of water - Dairy products (Dp), simply because dairy cattle and products (Dc and Dp) have rather similar patterns of linkages that those of dairy and beef cattle (Dc and Bc).[21]

[20]In case of gross output being the objective, i.e., when $\pi_i = 1$ for all i, the group output worth equals the sum of gross outputs of the extracted sectors and their indirect contributions to every other sectors' gross outputs.

[21]This can be proved formally using cluster analysis, which is, however, beyond the scope of the current paper. In this respect, our study has a link to Hoen (2002), who analyzes the groups of sectors with strong connections using different cluster identification methods and ends at choosing

The third observation from Table 1 is that sectors in the key group of size k are also part of the key group of size $k+1$, which raises a question of whether this is a general property or is a mere coincidence. It turns out that this is *not* true in general, i.e., the group target selection problem is not equivalent to a *sequential* key sector problem.[22] (The author can supply an interested reader by a hypothetical IO table that confirms the last statement.) One might (rightly) think that this fact is unfortunate from computational perspective, since this urges an analyst to compute the factor worths for *all* possible combinations of k from all n sectors, which, for instance, in our case with group of size 4 required to consider more than 13.6 million combinations, and that search process would be significantly reduced (i.e., to only 133 cases) if the key group problem and the sequential key sector problem would be equivalent.[23] Given that we have conjectured that the key group members are rather nonredundant and thus should be part of different clusters with similar linkage patterns, this result allows, at least theoretically, for "cluster switching" of sectors between clusters once the number of identified clusters changes. The phenomenon of "cluster switching" have been found, for example, in Howe and Stabler (1989). Hence, the fact that the key group problem requires to search for all possible combinations is, in fact, advantageous as it reveals cases of "cluster switching" if they do exist.

The forth observation is that a group of few industries accounts for the majority of the environmental factors, while generatioin of profits and salaries is relatively dispersed among sectors. So 58% and 53% of, respectively, water (direct and indirect) consumption and CO_2 emissions are due to the key groups of size 4 and 2 from the total of 136 sectors. The last technical observation is that the percentage decrease in overall factor usage/production upon extraction of groups is always

---

a *block diagonalization method* to suit best for clustering purpose. However, a word of caution is in place with respect to diagonalization method: it does *not* allow for "cluster switching". For instance, Howe and Stabler (1989) showed that an object may be assigned to totally different cluster if the number of identified clusters changes. In fact, this property of block diagonalization Hoen (2002) considers positively as other "cluster methods ... did not show this phenomenon [i.e., cluster switching] for sectors" (p. 139). However, the HEM allows for sector switching if one interprets the key group members in terms of different clusters' membership, at least theoretically (see the third observation in the text).

[22] By sequential (search) we mean the following: once the key group of size k has been identified, one needs only to add extra sector from all possible $n-k$ remaining industries in order to identify the key group of size $k+1$.

[23] This computation, however, took 2-3 minutes. The MATLAB program for finding key group can be provided by the author upon request.

smaller than the sum of the individual relative factor worths of sectors comprising the group, that is $\sum_{s=1}^{k} \omega_{i_s}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) > \omega_{i_1,i_2,\ldots,i_k}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi})$ for all $k = 2, 3, \ldots, n$ (see Appendix). This, however, does not have an economic meaning as we cannot add percentages of the relative factor worths, which do not sum up to 100% due to the fact that each sector's contribution is examined under the assumption that the rest of the sectors are active.

In order to compare the results of the generalized HEM to other indicators, in the last three columns of Table 1 we present the top 5 sectors with the largest factor multipliers, direct factor usage/generation, and factor responsibility. The first two indicators do not need explanation, hence we briefly discuss the third one. Multiplying the diagonalized matrix of the factor coefficients vector by the Leontief inverse gives the matrix $\hat{\boldsymbol{\pi}}\mathbf{B}$, whose ij-th element shows the amount of factor used/produced by sector i per unit final demand of sector j. Hence, the ij-th entry of the matrix $\hat{\boldsymbol{\pi}}\mathbf{B}\hat{\mathbf{f}}$ is the amount of the factor used/generated by sector i due to final demand of sector j, or equivalently, how much factor was consumed/produced by sector i *for* sector j. Thus, summing over all is gives the amount of the factor consumed/produced by all industries *for* sector j, which is the j-th element of the vector $\boldsymbol{\pi}'\mathbf{B}\hat{\mathbf{f}}$. In other words, this is the amount of the factor that sector j is responsible for, hence the term "responsibility".[24]

Multipliers are traditionally used to identify the importance of sectors. Table 1, however, shows that factor multipliers can give quite different results than those based on the HEM approach. This is expectable since factor worths besides the size of multipliers also take into account sectors' gross outputs size and net input dependencies. Rice (Ri) has the highest water use multiplier (7470 litres per A$ of its final demand), while it is not a member of the key groups of size $k \in [1, 4]$, and, moreover, it does not show up in the list of top 5 groups at all. Rice (Ri) though is the 5-th largest direct consumer of water (1.43 Tl), but it is not in the list of the top 5 responsible sectors. In case of CO_2 emissions, Forestry (Fr) has the largest CO_2 multiplier, but it is not a member of the key group of size $k < 3$. For gross operating surplus all four indicators give quite close outcomes with Ownership of dwellings (Dw) being the most important sector in all respect. Education (Ed) has the largest wages multiplier, and becomes a member of the key group of size 4. All in all, these

[24]See Hoen and Mulder (2003) for similar computation in analyzing the Dutch CO_2 emissions.

results do not mean that factor multipliers are useless from policy perspective. The advantage of multipliers lies in the price evaluation of commodities as multipliers are expressed per unit of final demand. In other words, industries with high factor multipliers are sensitive to changes in the factor price (see e.g., Dietzenbacher and Velázquez 2007). In our case, a pricing policy that tries to internalize the costs of using water and CO_2 emissions, would have the largest impact on the prices of, respectively, Rice (Ri) and Forestry (Fr).[25]

Notice also that for water use and CO_2 emissions there is a perfect correspondence in Table 1 between the key group members and the list of sectors with the largest direct factor usage/generation. But this is not always the case: the largest capacity of generating wages has Education (Ed, 14.6 Bln A$), which is not a member of the key group of size $k < 4$. Instead, Retail trade (Rt), which is *responsible* for the largest amount of wages (18.3 A$), is part of the key group of size $k \geq 2$. For water usage and CO_2 emissions dairy and meat products (Dp and Mp) are the most responsible sectors, while in both cases they do not show up as a part of the key groups. However, these industries are members of groups that are second in the list. All in all, it seems that the HEM approach takes into account both sectors' direct factor consumption/generation and sectors' responsibility in using/generating the factor by other industries. This is, of course, the specific advantage of using the generalized HEM, which fully considers all kinds of interlinkages associated with the hypothetically extracted sector(s).

5 Conclusion

In this paper we investigated the issue of identification of a key sector and a key group of sectors in the economy by a complete hypothetical extraction method (HEM). We show that for this purpose the analyst does not have perform the three step procedure of the HEM: delete the corresponding row(s) and column(s) of the input matrix, calculate the overall factor usage/production in the hypothetical case, and find the difference between the actual and hypothetical objectives. These steps

[25] In this respect for Australian case, Foran et al. (2005) regarding agricultural, forestry and food products state: "... the prices we pay for the products reflect the marginal cost of production, rather than the full resource and environmental costs of production. ... Moves to internalize the full costs of production in the final price of the market product may mean substantial price increases" (p. 1).

are rather excessive given that we have found quite simple formulas (measures of industries' factor worths) in getting the desired outcome.

We showed that the key sector problem and the key group problem have, in general, different solutions. This is demonstrated in the empirical application of the mentioned problems to the Australian economy for four factors of water use, CO_2 emissions, profits, and wages and salaries. In general, the key group has the highest group factor worth, which is directly related to the overall impact on aggregate factor usage/generation of an incremental changes in direct self-dependencies of the sectors comprising the group, and inversely related to own net input dependence of the group members. The last interpretation is a result of linking the HEM to the fields of influence method. The key sector/group problem can easily be used to address several policy issues simultaneously, for instance, finding key sectors in terms of increasing employment and decreasing emissions of greenhouse gases.

It is proved that an increase (resp. decrease) in an input coefficient never decreases (resp. increases) the factor worth of any sector. In both cases the necessary and sufficient condition for a strict change is that the sector supplying more per unit depends directly and/or indirectly on a sector whose worth is going to change.

We have added the expression "generalized" to the term HEM for two reasons. First, which is novel and the main contribution of this paper, a key sector search is extended formally to a key group search within the HEM framework, which also enables one to see their possible different outcomes. Second, these HEM problems are formally studied in terms of a general factor, not only gross output that used to be the main focus of the traditional HEM. Thus, depending on the research question, the general measures of industries' factor worths proposed in this paper will identify the corresponding key sector and/or key group of sectors.

References

Andreosso-O'Callaghan, B. and G. Yue: 2004, 'Intersectoral linkages and key sectors in China, 1987-1997'. *Asian Economic Journal* **18**, 165–183.

Ballester, C., A. Calvó-Armengol, and Y. Zenou: 2006, 'Who's who in networks. Wanted: the key player'. *Econometrica* **74**, 1403–1417.

Beynon, M. N. and M. Munday: 2008, 'Stochastic key sector analysis: an application to a regional input-output framework'. *Annals of Regional Science* **42**, 863–877.

Borgatti, S. P.: 2006, 'Identifying sets of key players in a social network'. *Computational and Mathematical Organization Theory* **12**, 21–34.

Burt, R. S.: 1992, *Structural Holes: The Social Structure of Competition*. Cambridge: Hardvard Univeristy Press.

Cai, J. and P. Leung: 2004, 'Linkage measures: a revisit and a suggested alternative'. *Economic System Research* **16**, 65–85.

Cardenete, M. A. and F. Sancho: 2006, 'Missing links in key sector analysis'. *Economic Systems Research* **18**, 319–325.

Cella, G.: 1984, 'The input-output measurement of interindustry linkages'. *Oxford Bulletin of Economics and Statistics* **46**, 73–84.

Centre for Integrated Sustainability Analysis, I. S. A.: 2005, 'Triple Bottom Line accounts of the Australian economy'. Unpublished data, The University of Sydney, Sydney, Australia.

Clements, B. J.: 1990, 'On the decomposition and normalization of interindustry linkages'. *Economics Letters* **33**, 337–340.

Daniels, P. L.: 1992, 'Barriers to sustainable development in natural resource-based economies: Australia as a case of study'. *Society and Natural Resources* **5**, 247–262.

Diamond, J.: 1975, 'Inter-industry indicators of employment potential'. *Applied Economics* **7**, 265–273.

Dietzenbacher, E.: 1992, 'The measurement of interindustry linkages: key sectors in the Netherlands'. *Economic Modelling* **9**, 419–437.

Dietzenbacher, E. and J. van der Linden: 1997, 'Sectoral and spartial linkages in the EC production structure'. *Journal of Regional Science* **37**, 235–257.

Dietzenbacher, E. and E. Velázquez: 2007, 'Analysing Andalusian virtual water trade in an input-output framework'. *Regional Studies* **41**, 185–196.

Duarte, R., J. Sánchez-Chóliz, and J. Bielsa: 2004, 'Water use in the Spanish economy: an input-output approach'. *Ecological Economics* **43**, 71–85.

Everett, M. G. and S. P. Borgatti: 1999, 'The centrality of groups and classes'. *Journal of Mathematical Sociology* **23**, 181–201.

Everett, M. G. and S. P. Borgatti: 2005, 'Extending centrality'. In: S. J. Carrington, Peter J. and W. Stanley (eds.): *Models and Methods in Social Network Analysis*. Cambridge: Cambridge University Press, pp. 57–76.

Foran, B., M. Lenzen, and C. Dey: 2005, *Balancing Act - A Triple Bottom Line Account of the Australian Economy*. Internet site http://www.isa.org.usyd.edu.au, Canberra, ACT, Australia, CSIRO Resource Futures and The University of Sydney.

Fritz, O., M. Sonis, and G. J. Hewings: 2002, 'Direct and indirect industrial pollution generation: a field of influence approach'. In: G. J. Hewings, M. Sonis, and D. Boyce (eds.): *Trade, Networks and Hierarchies. Modeling Regional and Interregional Economies*. Berlin: Spriger-Verlag, pp. 433–443.

Gould, B. W. and S. N. Kulshreshtha: 1986, 'An interindustry analysis of structural change and energy use linkages in the Saskatchewan economy'. *Energy Economics* **8**, 186–196.

Groenewold, N., A. Hagger, and J. R. Madden: 1987, 'The measurement of industry employment contribution in an input-output model'. *Regional Studies* **21**, 255–263.

Groenewold, N., A. Hagger, and J. R. Madden: 1993, 'Measuring industry importance: an Australian application'. *Annals of Regional Science* **27**, 175–182.

Heimler, A.: 1991, 'Linkages and vertical integration in the Chinese economy'. *Review of Economics and Statistics* **73**, 261–267.

Henderson, H. V. and S. R. Searle: 1981, 'On deriving the inverse of a sum of matrices'. *SIAM Review* **23**, 53–60.

Henriques, A. and J. Richardson (eds.): 2004, *The Triple Bottom Line: Does it All Add Up?* London: Earthscan.

Hewings, G. J., M. Fonseca, J. J. Guilhoto, and M. Sonis: 1989, 'Key sectors and structural change in the Brazilian economy: a comparison of alternative approaches and their policy implications'. *Jounal of Policy Modeling* **11**, 67–90.

Hirschman, A.: 1958, *The Strategy of Economic Development.* New Haven: Yale University Press.

Hoen, A. and M. Mulder: 2003, 'Explaining Dutch emissions of CO_2; a decomposition analysis'. *CPB Discussion Paper* **24**.

Hoen, A. R.: 2002, 'Indentifying linkages with a cluster-based methodology'. *Economic Systems Research* **14**, 131–146.

Howe, E. C. and J. C. Stabler: 1989, 'Canada divided: the optimal division of an economy into regions'. *Journal of Regional Science* **29**, 191–211.

Jones, L. P.: 1976, 'The measurement of Hirshmanian linkages'. *Quarterly Journal of Economics* **90**, 323–333.

Kol, J.: 1991, 'Key sectors, comparative advantage, and international shifts in employment: a case study for Indonesia, South Korea, Mexico and Pakistan and their trade relations with the European Community'. In: W. Peterson (ed.): *Advances in Input-Output Analysis.* New York: Oxford University Press, pp. 199–210.

Lenzen, M.: 2003, 'Environmentally important paths, linkages and key sectors in the Australian economy'. *Structural Change and Economic Dynamics* **14**, 1–34.

Los, B.: 2004, 'Identification of strategic industries: a dynamic perspective'. *Papers in Regional Science* **83**, 669–698.

Meller, P. and M. Marfán: 1981, 'Small and large industry: employment generation, linkages and key sectors'. *Economic Development and Cultural Change* **29**, 263–274.

Midmore, P., M. Munday, and A. Roberts: 2006, 'Assessing industry linkages using regional input-output tables'. *Regional Studies* **40**, 329–343.

Miller, R. E. and P. D. Blair: 1985, *Input-Output Analysis: Foundations and Extensions*. Englewoods Cliffs, New Jersey: Prentice-Hall.

Miller, R. E. and M. L. Lahr: 2001, 'A taxonomy of extractions'. In: M. L. Lahr and R. E. Miller (eds.): *Regional Science Perspectives in Economics: A Festschrift in Memory of Benjamin H. Stevens*. Amsterdam: Elsevier Science, pp. 407–441.

Paelinck, J., J. de Caevel, and D. J.: 1965, 'Analyse Quantitative de Certaines Phénomènes du Développment Régional Polarisé: Essai de Simulation Statique d'itérarires de Propogation'. In: *Problémes de Conversion Éconmique: Analyses Théoretiques et Études Appliquées*. Paris, pp. 341–387, M.-Th. Génin.

Porter, M. E.: 1990, *Competitive Advantage of Nations*. New York, NY: Free Press.

Rasmussen, P. N.: 1956, *Studies in Inter-Sectoral Relations*. Amsterdam: North-Holland.

Roberts, B. M.: 1995, 'Structural change in Poland, 1980-90: evidence from social accounting multipliers and linkage analysis'. *Applied Economics* **7**, 291–308.

Schultz, S.: 1977, 'Approaches to identifying key sectors empirically by means of input-output analysis'. *Journal of Development Studies* **14**, 77–96.

Sherman, J. and W. J. Morrison: 1950, 'Adjustment of an inverse matrix corresponding to a change in one element of a given matrix'. *Annals of Mathematical Statistics* **21**, 124–127.

Song, Y. and C. Liu: 2007, 'An input-output approach for measuring real estate sector linkages'. *Journal of Property Research* **24**, 71–91.

Song, Y., C. Liu, and G. Langston: 2006, 'Linkage measures of the construction sector using the hypothetical extraction method'. *Construction Management and Economics* **24**, 579–589.

Sonis, M., J. J. Guilhoto, G. J. Hewings, and E. B. Martins: 1995, 'Linkages, key sectors, and structural change: some new perspectives'. *Developing Economies* **33**, 233–270.

Sonis, M. and G. J. Hewings: 1989, 'Error and sensitivity input-output analysis: a new approach'. In: R. E. Miller, K. R. Polenske, and A. Z. Rose (eds.): *Frontiers of Input-Output Analysis*. New York: Oxford University Press, pp. 232–244.

Sonis, M. and G. J. Hewings: 1992, 'Coefficient change in input-output models: theory and applications'. *Economic Systems Research* **4**, 143–157.

Strassert, G.: 1968, 'Zur bestimmung strategischer sektoren mit hilfe von von input-output modellen'. *Jahrbűcher fűr Nationalőkonomie und Statistik* **182**, 211–215.

Temurshoev, U.: 2008, 'Who's who in networks. Wanted: the key group'. *Net Institute Working Paper* **08-08**.

Temurshoev, U.: 2009, 'A connection: network games, coalitional game, and hypothetical extraction'. University of Groningen, Mimeo.

Ten Raa, T.: 2005, *The Economics of Input-Output Analysis*. Cambridge: Cambridge Univeristy Press.

Yotopoulos, P. A. and J. B. Nugent: 1973, 'A balanced-growth version of the linkage hypothesis: a test'. *Quarterly Journal of Economics* **87**, 157–171.

Zeng, L.: 2001, 'A property of the Leontief inverse and its applications to comparative static analysis'. *Economic Systems Research* **13**, 299–315.

Appendix

Proof of Lemma 1. We give our proof within the IO framework. First note that the matrix $\mathbf{I} - \mathbf{e}_i\mathbf{e}_i'$ has ones in all diagonal entries except for its ii-th position, and zero otherwise. Hence, $\mathbf{A}^{-i} = (\mathbf{I} - \mathbf{e}_i\mathbf{e}_i')\mathbf{A}(\mathbf{I} - \mathbf{e}_i\mathbf{e}_i')$. We make use of the well-known formula of the inverse of a sum of matrices (see e.g., Henderson and Searle 1981):

$$(\mathbf{X} - \mathbf{U}\mathbf{D}^{-1}\mathbf{Z})^{-1} = \mathbf{X}^{-1} + \mathbf{X}^{-1}\mathbf{U}(\mathbf{D} - \mathbf{Z}\mathbf{X}^{-1}\mathbf{U})^{-1}\mathbf{Z}\mathbf{X}^{-1}, \tag{A1}$$

$$(\mathbf{X} + \mathbf{u}\mathbf{z}')^{-1} = \mathbf{X}^{-1} - \frac{1}{1 + \mathbf{z}'\mathbf{X}^{-1}\mathbf{u}}\mathbf{X}^{-1}\mathbf{u}\mathbf{z}'\mathbf{X}^{-1}. \tag{A2}$$

Since $\mathbf{e}_i\mathbf{e}_i'\mathbf{e}_i\mathbf{e}_i' = \mathbf{e}_i\mathbf{e}_i'$ (as $\mathbf{e}_i'\mathbf{e}_i = 1$), one can easily confirm that $(\mathbf{I} - \mathbf{e}_i\mathbf{e}_i')(\mathbf{I} - \mathbf{e}_i\mathbf{e}_i') = \mathbf{I} - \mathbf{e}_i\mathbf{e}_i'$. Then using (A1) it follows that

$$\mathbf{B}^{-i} = (\mathbf{I} - (\mathbf{I} - \mathbf{e}_i\mathbf{e}_i')\mathbf{A}(\mathbf{I} - \mathbf{e}_i\mathbf{e}_i'))^{-1} = \mathbf{I} + (\mathbf{I} - \mathbf{e}_i\mathbf{e}_i')[\mathbf{A}^{-1} - (\mathbf{I} - \mathbf{e}_i\mathbf{e}_i')]^{-1}(\mathbf{I} - \mathbf{e}_i\mathbf{e}_i'). \tag{A3}$$

Using (A1) again we can write the Leontief inverse as $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + (\mathbf{A}^{-1} - \mathbf{I})^{-1}$, which implies that $(\mathbf{A}^{-1} - \mathbf{I})^{-1} = \mathbf{B} - \mathbf{I}$. This together with (A2) allows us to write the inverse in the right-hand side of (A3) as:

$$((\mathbf{A}^{-1} - \mathbf{I}) + \mathbf{e}_i\mathbf{e}_i')^{-1} = \mathbf{B} - \mathbf{I} - \frac{1}{b_{ii}}(\mathbf{B} - \mathbf{I})\mathbf{e}_i\mathbf{e}_i'(\mathbf{B} - \mathbf{I}), \tag{A4}$$

where the last follows from the fact that $\mathbf{e}_i'(\mathbf{B} - \mathbf{I})\mathbf{e}_i = b_{ii} - 1$. Plugging (A4) in (A3) and using the fact that $\mathbf{e}_i'\mathbf{B}\mathbf{e}_i = b_{ii}$, some simple matrix multiplication yields $\mathbf{B}^{-i} = \mathbf{e}_i\mathbf{e}_i' + \mathbf{B} - \frac{1}{b_{ii}}\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B}$, which completes the proof. □

Derivation of problem (3). The objective function in problem (2) is $\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-i} = \boldsymbol{\pi}'(\mathbf{B}\mathbf{f} - \mathbf{B}^{-i}\mathbf{f}^{-i})$. Adding and subtracting $\mathbf{B}^{-i}\mathbf{f}$ to the expression in the brackets gives $\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-i} = \boldsymbol{\pi}'(\mathbf{B} - \mathbf{B}^{-i})\mathbf{f} + \boldsymbol{\pi}'\mathbf{B}^{-i}(\mathbf{f} - \mathbf{f}^{-i})$. It is apparent that $\mathbf{f} - \mathbf{f}^{-i} = f_i\mathbf{e}_i$. This together with Lemma 1 yields

$$\begin{aligned}
\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-i} &= \boldsymbol{\pi}'\left(\frac{1}{b_{ii}}\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B} - \mathbf{e}_i\mathbf{e}_i'\right)\mathbf{f} + f_i\boldsymbol{\pi}'\left(\mathbf{B} - \frac{1}{b_{ii}}\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B} + \mathbf{e}_i\mathbf{e}_i'\right)\mathbf{e}_i \\
&= \frac{1}{b_{ii}}\boldsymbol{\pi}'\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B}\mathbf{f} - f_i\pi_i + f_i\boldsymbol{\pi}'\mathbf{B}\mathbf{e}_i - \frac{f_i}{b_{ii}}\boldsymbol{\pi}'\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B}\mathbf{e}_i + f_i\pi_i \\
&= \frac{1}{b_{ii}}\boldsymbol{\pi}'\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B}\mathbf{f} + f_i\boldsymbol{\pi}'\mathbf{B}\mathbf{e}_i - \frac{f_i}{b_{ii}}\boldsymbol{\pi}'\mathbf{B}\mathbf{e}_i\mathbf{e}_i'\mathbf{B}\mathbf{e}_i = \frac{1}{b_{ii}}\mathbf{m}_\pi'\mathbf{e}_i\mathbf{e}_i'\mathbf{x},
\end{aligned}$$

where the last term follows since $\mathbf{e}_i'\mathbf{B}\mathbf{e}_i = b_{ii}$. □

Proof of Theorem 2. Using the definitions of the factor worth, factor multiplier and equation (1), we have $\omega_i^\pi(\mathbf{A},\mathbf{f},\boldsymbol{\pi}) = \frac{1}{b_{ii}} m_i^\pi x_i = \left(\sum_{j=1}^n \pi_j b_{ji}\right) \sum_{j=1}^n \frac{b_{ij}}{b_{ii}} f_j$. Then,

$$\Delta_i^\pi \equiv \omega_i^\pi(\widetilde{\mathbf{A}},\mathbf{f},\boldsymbol{\pi}) - \omega_i^\pi(\mathbf{A},\mathbf{f},\boldsymbol{\pi}) = \left(\sum_{j=1}^n \pi_j \tilde{b}_{ji}\right) \sum_{j=1}^n \frac{\tilde{b}_{ij}}{\tilde{b}_{ii}} f_j - \left(\sum_{j=1}^n \pi_j b_{ji}\right) \sum_{j=1}^n \frac{b_{ij}}{b_{ii}} f_j,$$

where $\tilde{b}_{ij}$ is a generic element of $\widetilde{\mathbf{B}} = (\mathbf{I}-\widetilde{\mathbf{A}})^{-1}$. Adding and subtracting $\left(\sum_j \pi_j \tilde{b}_{ji}\right) \sum_j \frac{b_{ij}}{b_{ii}} f_j$ to the last expression yields

$$\Delta_i^\pi = \left(\sum_{j=1}^n \pi_j \tilde{b}_{ji}\right) \sum_{j=1}^n \left(\frac{\tilde{b}_{ij}}{\tilde{b}_{ii}} - \frac{b_{ij}}{b_{ii}}\right) f_j + \left(\sum_{j=1}^n \pi_j (\tilde{b}_{ji} - b_{ji})\right) \sum_{j=1}^n \frac{b_{ij}}{b_{ii}} f_j. \qquad \text{(A5)}$$

From (3) (see also fn. 12) it follows that $\tilde{b}_{ij} = b_{ij} + \epsilon_i b_{cj}$, where $\epsilon_i = \alpha b_{ir}/(1-\alpha b_{cr})$. Therefore,

$$\frac{\tilde{b}_{ij}}{\tilde{b}_{ii}} - \frac{b_{ij}}{b_{ii}} = \frac{b_{ij} + \epsilon_i b_{cj}}{b_{ii} + \epsilon_i b_{ci}} - \frac{b_{ij}}{b_{ii}} = \frac{\epsilon_i (b_{ii} b_{cj} - b_{ci} b_{ij})}{b_{ii}(b_{ii} + \epsilon_i b_{ci})}.$$

Plugging the last expression in (A5) and using $\tilde{b}_{ij} = b_{ij} + \epsilon_i b_{cj}$ gives

$$\Delta_i^\pi = \epsilon_i \left[\left(\sum_{j=1}^n \pi_j \tilde{b}_{ji}\right) \sum_{j=1}^n \left(\frac{b_{ii} b_{cj} - b_{ci} b_{ij}}{b_{ii}(b_{ii} + \epsilon_i b_{ci})}\right) f_j + \left(b_{ci} \sum_{j=1}^n \pi_j\right) \sum_{j=1}^n \frac{b_{ij}}{b_{ii}} f_j \right]. \qquad \text{(A6)}$$

One of the well-know property of the Leontief inverse is that $b_{ii} \geq 1$ and $b_{ii} > b_{ij} \geq 0$ for all i and all $j \neq i$. Theorem 1 in Zeng (2001) shows that $b_{ii} b_{cj} \geq b_{ci} b_{ij}$, with strict inequality holding when $j = c \neq i$. Hence, $\sum_j (b_{ii} b_{cj} - b_{ci} b_{ij}) f_j > 0$ for all $i \neq c$ (assuming that $f_j > 0$ for all j). It is not difficult to see that for $i = c$ every term in this sum is zero, hence the first term of Δ_c^π in (A6) (when $i = c$) vanishes, however, its second term is positive as $b_{cc} \geq 1$ with the two other sums being positive. So the expression within the square brackets in (A6) is always positive, hence the sign of Δ_i^π will depend only on ϵ_i, which is zero whenever $b_{ir} = 0$, and otherwise positive if $\alpha > 0$, while negative if $\alpha < 0$. This completes the proof. □

Proof of Lemma 2. Lemma 1 in Temurshoev (2008) in the framework of social network analysis is mathematically equivalent to Lemma 2 in this paper. Hence, see the proof of Lemma 1 in Temurshoev (2008). □

Derivation of problem (5). As in derivation of problem (3), the objective function in problem (4) can be rewritten as $\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-\{i_1,\ldots,i_k\}} = \boldsymbol{\pi}'(\mathbf{B} - \mathbf{B}^{-\{i_1,\ldots,i_k\}})\mathbf{f} + \boldsymbol{\pi}'\mathbf{B}^{-\{i_1,\ldots,i_k\}}(\mathbf{f} - \mathbf{f}^{-\{i_1,\ldots,i_k\}})$, where $\mathbf{f} - \mathbf{f}^{-\{i_1,\ldots,i_k\}} = \sum_{s=1}^k f_{i_s} \mathbf{e}_{i_s} = \mathbf{E}\mathbf{E}'\mathbf{f}$. This together with Lemma 2 and the fact that $\mathbf{E}\mathbf{E}'\mathbf{E}\mathbf{E}' = \mathbf{E}\mathbf{E}'$ gives

$$\begin{aligned}
&\boldsymbol{\pi}'\mathbf{x} - \boldsymbol{\pi}'\mathbf{x}^{-\{i_1,\ldots,i_k\}} \\
&\quad = \boldsymbol{\pi}' \left[\mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{B} - \mathbf{E}\mathbf{E}'\right] \mathbf{f} + \boldsymbol{\pi}' \left[\mathbf{B} - \mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{B} + \mathbf{E}\mathbf{E}'\right] \mathbf{E}\mathbf{E}'\mathbf{f} \\
&\quad = \boldsymbol{\pi}'\mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{B}\mathbf{f} - \mathbf{E}\mathbf{E}'\mathbf{f} + \boldsymbol{\pi}'\mathbf{B}\mathbf{E}\mathbf{E}'\mathbf{f} - \boldsymbol{\pi}'\mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{B}\mathbf{E}\mathbf{E}'\mathbf{f} + \mathbf{E}\mathbf{E}'\mathbf{f} \\
&\quad = \boldsymbol{\pi}'\mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{B}\mathbf{f} + \boldsymbol{\pi}'\mathbf{B}\mathbf{E}\mathbf{E}'\mathbf{f} - \boldsymbol{\pi}'\mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{B}_{kk}\mathbf{E}'\mathbf{f} = \boldsymbol{\pi}'\mathbf{B}\mathbf{E}\mathbf{B}_{kk}^{-1}\mathbf{E}'\mathbf{B}\mathbf{f},
\end{aligned}$$

which is exactly the objective of the key group problem (5). □

Proof of the inequality $\sum_{s=1}^{k} \omega_{i_s}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) > \omega_{i_1,i_2,\ldots,i_k}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi})$. It is easy to show that $\sum_{s=1}^{k} \frac{1}{b_{i_s i_s}} \mathbf{e}_{i_s} \mathbf{e}_{i_s}' = \mathbf{E}\mathbf{E}'\widehat{\mathbf{B}}^{-1}\mathbf{E}\mathbf{E}'$, where $\widehat{\mathbf{B}}$ is a diagonal matrix with b_{ii} on its main diagonal and zeros elsewhere. Define $\widehat{\mathbf{B}}_{kk}^{-1} \equiv \mathbf{E}'\widehat{\mathbf{B}}^{-1}\mathbf{E}$ as the reduced diagonal matrix with $1/b_{i_s i_s}$ on its main diagonal with $s = 1, \ldots, k$. Thus, the sum of the individual factor worths of the k industries is $\sum_{s=1}^{k} \omega_{i_s}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) = \sum_{s=1}^{k} \frac{1}{b_{i_s i_s}} \mathbf{m}_{\pi}' \mathbf{e}_{i_s} \mathbf{e}_{i_s}' \mathbf{x} = \mathbf{m}_{\pi}' \left(\sum_{s=1}^{k} \frac{1}{b_{i_s i_s}} \mathbf{e}_{i_s} \mathbf{e}_{i_s}' \right) \mathbf{x} = \mathbf{m}_{\pi}' \mathbf{E} \widehat{\mathbf{B}}_{kk}^{-1} \mathbf{E}' \mathbf{x}$. Using the definition of the group factor worth, then implies

$$\begin{aligned} \sum_{s=1}^{k} \omega_{i_s}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) - \omega_{i_1,i_2,\ldots,i_k}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) &= \mathbf{m}_{\pi}' \mathbf{E} \widehat{\mathbf{B}}_{kk}^{-1} \mathbf{E}' \mathbf{x} - \mathbf{m}_{\pi}' \mathbf{E} \mathbf{B}_{kk}^{-1} \mathbf{E}' \mathbf{x} \\ &= \mathbf{m}_{\pi}' \mathbf{E} \left[\widehat{\mathbf{B}}_{kk}^{-1} - \mathbf{B}_{kk}^{-1} \right] \mathbf{E}' \mathbf{x}. \end{aligned}$$

Suppose the above difference is negative. Then using the fact that the group factor worth is always positive, we will have $\omega_{i_1,i_2,\ldots,i_k}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) > \mathbf{m}_{\pi}' \mathbf{E} \left[\widehat{\mathbf{B}}_{kk}^{-1} - \mathbf{B}_{kk}^{-1} \right] \mathbf{E}' \mathbf{x}$, which holds if and only if $\mathbf{B}_{kk}^{-1} > \widehat{\mathbf{B}}_{kk}^{-1} - \mathbf{B}_{kk}^{-1}$, or $2\mathbf{B}_{kk}^{-1} > \widehat{\mathbf{B}}_{kk}^{-1}$. The last condition, however, is not true given the fact that due to Leontief matrix properties there will be always negative off-diagonal elements in the matrix $\mathbf{B}_{kk}^{-1}$. Hence, by contradiction, we have proved that $\sum_{s=1}^{k} \omega_{i_s}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi}) \geq \omega_{i_1,i_2,\ldots,i_k}^{\pi}(\mathbf{A}, \mathbf{f}, \boldsymbol{\pi})$. It is easy to see that the equality will hold only if $\widehat{\mathbf{B}}_{kk}^{-1} = \mathbf{B}_{kk}^{-1}$, which holds only when $\mathbf{A} = \mathbf{O}$, where $\mathbf{O}$ is a zero matrix. However, the last condition is nonsense given that there are always sectors that are somehow interrelated (i.e. $\mathbf{A} > \mathbf{O}$). $\square$

Table 2: Codes assigned to 136 Australian sectors

Sym.	Industry	Sym.	Industry
Ac	Insecticides, pesticides and other agricultural chemicals	Lm	Lime
Ai	Aircraft	Lp	Leather and leather products
Al	Aluminium alloys and aluminium recovery	Ma	Agricultural, mining and construction machinery
Ao	Alumina	Mi	Mineral and glass wool and other non-metallic mineral products
Ap	Automotive petrol	Mn	Exploration and services to mining
At	Air and space transport	Mp	Meat and meat products
Ba	Barley, unmilled	Ms	Legal, accounting, marketing and business management services
Bc	Beef cattle	Mv	Motor vehicles and parts, other transport equipment
Bk	Banking	Nb	Non-residential buildings, roads, bridges and other construction
Bl	Black coal	Ne	Newspapers, books, recorded media and other publishing
Bm	Beer and malt	Nf	Non-ferrous metal recovery and basic products
Bp	Bread, cakes, biscuits and other bakery products	Ng	Natural gas
Br	Brown coal, lignite	Oc	Adhesives, inks, polishes and other chemical products
Bs	Typing, copying, staff placement and other business services	Oe	Photographic, optical, medical and radio equipment, watches
Bt	Bus and tramway transport services	Of	Oils and fats
Bu	Prefabricated buildings	Oi	Crude oil
Bv	Soft drinks, cordials and syrups	Om	Coins, jewellery, sporting goods and other manufacturing
Bx	Bauxite	Os	Police, interest groups, fire brigade and other services
Cc	Concrete and mortar	Ot	Cable car, chair lift, monorail and over-snow transport
Ce	Cement	Pa	Paper containers and products
Cg	Services to agriculture, ginned cotton, shearing and hunting	Pc	Petroleum bitumen, refinery LPG and other refinery products
Ch	Basic chemicals	Pd	Property developer, real estate and other property services
Cl	Clothing	Pe	Poultry and eggs
Cm	Communication services	Pg	Pigs
Cn	Confectionery	Ph	Pharmaceutical goods for human use
Co	Copper	Pi	Pipeline transport services
Cp	Plaster and other concrete products	Pl	Plastic products
Cr	Bricks and other ceramic products	Pp	Pulp, paper and paperboard
Cs	Childminding and other community care services	Pr	Printing, stationery and services to printing
Ct	Cosmetics and toiletry preparations	Ps	Hairdressing, goods hiring, laundry and other personal services
Cu	Libraries, parks, museums and the arts	Pt	Paints
Dc	Dairy cattle and untreated whole milk	Rb	Residential building, construction, repair and maintenance
De	Soap and other detergents	Rd	Road freight transport services
Df	Defence	Rf	Railway freight transport services
Dp	Dairy products	Rh	Repairs of household and business equipment
Dw	Ownership of dwellings	Ri	Rice, in the husk
Ed	Education	Rp	Railway passenger transport services
Ee	Cable, wire, batteries, lights and other electrical equipment	Rs	Sport, gambling and recreational services
El	Electricity supply	Rt	Retail trade
En	Electronic equipment, photocopying, gaming machines	Ru	Rubber products
Eq	Pumps, bearings, air conditioning and other equipment	Rv	Repairs of motor vehicles, agricultural and other machinery
Et	Motion picture, radio and television services	Rw	Railway equipment
Fc	Flour, cereal foods, rice, pasta and other flour mill products	Sb	Ships and boats
Fd	Raw sugar, animal feeds, seafoods, coffee and other foods	Sc	Seed cotton
Fe	Mixed fertilisers	Sf	Security broking and dealing and other services to finance
Fi	Commercial fishing	Sg	Sand, gravel and other construction materials mining
Fm	Nuts, bolts, tools and other fabricated metal products	Sh	Sheet containers and other sheet metal products
Fn	Money market corporation and other non-bank finance	Si	Financial asset investors and holding company services
Fo	Gas oil, fuel oil	Sm	Frames, mesh and other structural metal products
Fp	Vegetables, fruit, juices and other fruit and vegetable products	Sp	Water transport
Fr	Forestry and services to forestry	St	Travel agencies, forwarding and other services to transport
Fu	Furniture	Su	Sugar cane
Fw	Footwear	Sw	Softwoods, conifers
Ga	Gas production and distribution	Sz	Silver and zinc ores
Gd	Sanitary and garbage disposal services	Ta	Taxi and hired car with driver
Gl	Gold and lead	Ti	Sawn timer, woodchips and other sawmill products
Gp	Glass and glass products	To	Tobacco products
Gv	Government administration	Tp	Carpets, curtains, tarpaulins, sails, tents and other textiles
Hh	Household appliances and hot water systems	Ts	Scientific research, technical and computer services
Ho	Accommodation, cafes and restaurants	Tx	Processed wool, textile fibres, yarns and woven fabrics
Hs	Health services	Uo	Uranium, nickel, tin, manganese and other non-ferrous metal ores
Hw	Hardwoods, brushwoods, scrubwoods, hewn and other timber	Vf	Vegetable and fruit growing, hay, plant nurseries, flowers
In	Insurance	Wa	Water supply, sewerage and drainage services
Io	Iron ores	Wh	Wheat, legumes for grain, oilseeds, oats and other grains
Is	Basic iron and steel, pipes, tubes, sheets, rods, bars, rails, fittings	Wo	Sheep and shorn wool
Ke	Kerosene and aviation jet fuel	Wp	Plywood, window frames, doors and other wood products
Kn	Knitting mill products	Ws	Wine and spirits
Lg	Liquefied natural gas, liquefied natural petrol	Wt	Wholesale trade