15th International Conference on Input-Output Techniques
Beijing, China
June 27 to July 1, 2005

Structural Decomposition Analysis of Air Emissions in Denmark 1980-2002

Peter Rørmose

and

Thomas Olsen

Statistics Denmark

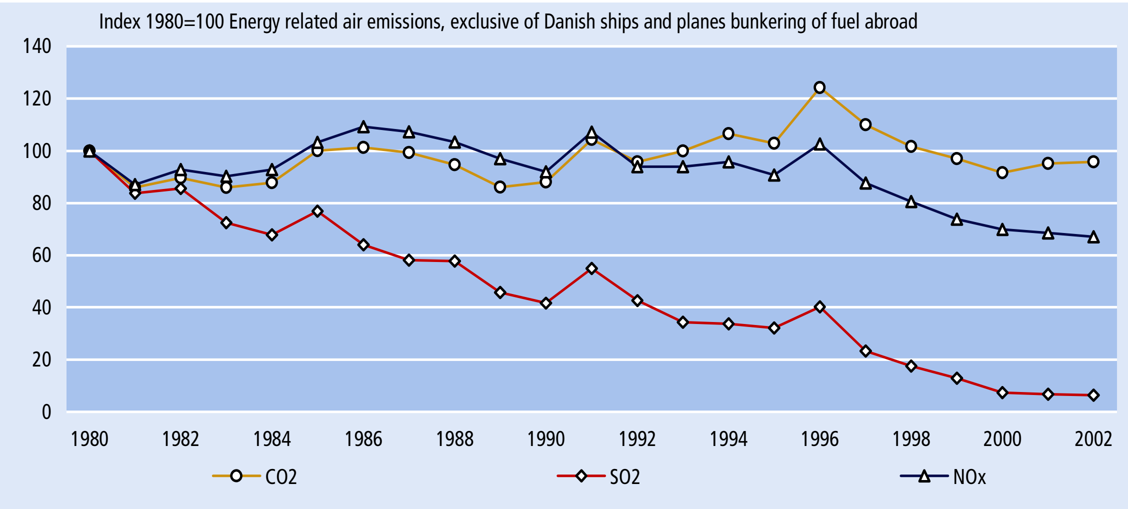
Contents

1. Introduction

This paper[1] presents analyses of air-emissions related to the use of energy in Denmark 1980-2002. It is based on the newly constructed time series 1980 – 2002 of Danish CO_2, SO_2 and NO_x air emissions. The time series, which initially will be introduced briefly, is an integrated part of the Danish NAMEA (National Accounting Matrix including Environmental Accounts)[2] accounts.

The new time series replace the time series for 1980-1992, which was based on the classification of the "old" (before the SNA 95 revision) national accounts (described in Jensen & Pedersen, 1998), and the time series for 1990-1999, which follows the existing national accounts. The new emissions accounts include the most recent information on emissions factors from the Danish CORINAIR database from the Danish National Environmental Research Institute. This up-to-date information has formed the basis for the estimation of the entire time series in order to ensure consistency and comparability over time.

Figure 1.1 Air Emissions related to use of energy 1980 - 2002



The new data show that the energy related emissions of CO_2 in Denmark is the same in 2002 as it was in 1980. In the period between we have been at 85% of the 1980 level as well as 120%. The emissions of NO_x show an almost similar development, however with a more substantial 35% decrease in recent years. The big winner is SO_2 that is now at about only 5% of the level in 1980. Thus, with the time series a basis for analyzing and modeling the trends in the air emissions - especially the longer-term developments - exists. In the case of this paper we will focus on the reasons for the development in the series above.

A huge part of man-made emissions of CO_2, NO_x and SO_2 is related to the combustion of energy. The combustion takes place as a response to the demand for energy, which is dependent of the size and structure of the economy, and is determined in an interaction between the various sectors on the basis of prices, legislation and so on. Together with the many technical possibilities for producing and distributing energy it forms a complex chain of different driving forces behind the emissions to air. In order to get a good understanding of the historical changes in the emissions as a tool in the process of planning a more sustainable

[1] The work on analysis of changes in air emissions in Denmark has benefited from a grant from the Commission of the European Communities (DG Eurostat/B1 Grant agreement nr. 200141200007).

[2] The conceptual framework for the NAMEA air emissions accounts is described in detail in NAMEA for Air Emissions - Compilation Guide (EUROSTAT, 2003).

economy, it can be very useful to be able to separate these driving forces into individual components. For such purpose decomposition analysis is a strong tool that can reveal the underlying factors. In the paper it is shown how the NAMEA air emissions accounts can be used for decomposition analysis of the development in air emissions.

Decomposition analysis is a way to ascribe the change in a variable of interest to the sum of changes in a number of other variables. Following the description of the input-output based techniques of decomposition analysis, a set of specific Danish decomposition models is presented. The 1980-2002 time series of emissions and other energy related matrices and vectors are combined with the corresponding Danish (130 x 130 industry) input-output tables and final demand tables in order to reveal to causes of the development we see in figure 1.1.

2. Air Emission Accounts in Denmark, NAMEA

2.1. Environmental Accounts

The Danish environmental accounts are constructed as satellite accounts linked with the national accounts and the input-output tables. The basic flow accounts are formed as supply and use tables as described in SEEA (2003).

2.1.1. Energy Accounts

Statistics Denmark collects and maintains quite large annual databases of energy use organized in the energy accounts. Here input of various energy types are balanced with the use of energy. The collection of these data is closely connected with compilation of the national accounts in Denmark. The accounts are organised in such a way that they are directly compatible with the national accounts at the most detailed industry level. They describe the supply and use of energy and keep account of 40 different energy carriers such as oil, gas, coal, gasoline and wood, straw and wind power. The energy flow accounts are broken down by 130 industries as well as various types of final demand (private consumption, exports etc.). This ensures consistency with the national accounts with regard to classification of activities.

Monetary accounts are, as already mentioned, constructed complementary to the physical accounts for energy. These are fully consistent with the physical accounts and they are used as basis for the national accounts as far as the description of the monetary transactions related to energy is concerned. In this way, full consistency between the environmental accounts and the national accounts is ensured.

2.1.2. Air emission Accounts

The air emissions accounts, which are made up in tonnes, distinguish between eight substances (CO_2, SO_2, NO_x, CH_4, N_2O, NMVOC, NH_3, and CO). For energy related emissions, the accounts include a breakdown of the emissions by the same 40 types of energy, which are included in the energy accounts. Furthermore, all information on emissions is broken down by 130 industries and households.

2.2. The Energy Accounts: Data sources and method

The supply side of the energy accounts is determined by the commodity statistics and the external trade statistics both of which are made up in physical as well as monetary values.

The use side of the energy accounts relies on information on the energy sector from the Danish Energy Agency, census' of the energy consumption in the industries and data on reimbursement of energy taxes and data on employment.

Ships and planes belonging to Danish shipping and airline companies take in a huge amount of fuel oil in foreign harbours and jet petroleum in foreign air ports. The expenditures connected to this bunkering are included in the Danish national accounts together with other Danish expenditures abroad.

In order to give a complete picture of the energy consumption related to the Danish activity, the bunkering by Danish ships and planes in foreign countries is included in the energy accounts. The amount of fuel oil bunkered has been estimated from the expenses in DKK for bunkering by Danish companies and corresponding fuel oil and jet petroleum prices.

The measurement of supply and use of energy can be based on the so-called "direct energy method" as well as the "gross energy method". According to the direct energy method, the full consumption of energy carriers should be reckoned among those who actually use it - primarily the conversion sector. This means that the power plants and the district heating facilities will be the absolute main polluters. The direct energy method is the basis for the analyses in this paper.

To an increasing degree electricity is traded across the Danish borders. Denmark imports electricity mainly from Norway and Sweden. For reasons of simplicity it has traditionally been assumed that this electricity is produced with the same technology as if it had been produced in Denmark and thus, has the same pollution consequences as the Danish electricity production has. This will also be the attitude in this paper, although this is a very erroneous assumption. Because of a lot of emission-free electricity production in Sweden and Norway, the emission consequences of the imported electricity are overstated.

The table below shows an aggregated version of the Danish energy accounts.

Table 2.1 Direct energy consumption and energy intensities 1980 and 2002*, 1995-prices

		1980	2002*	1980	2002*
		TJ		TJ/mill. DKK	
1	Agriculture, fishing and quarrying	48 784	79 223	0.87	0.91
2	Manufacturing	434 574	510 889	1.30	1.03
3	Electricity, gas and water supply	343 780	363 195	17.56	10.20
4	Construction	11 104	15 963	0.10	0.11
5	Wholesale and retail trade; hotels, restaurants	44 621	42 730	0.25	0.15
6	Transport, storage and communication	159 314	337 127	1.72	1.47
7	Financial intermediation, business activities	9 347	15 753	0.05	0.04
8	Public and personal services	38 132	41 109	0.15	0.11
	Total industries	**1 089 655**	**1 405 989**	**0.87**	**0.68**
	Households	247 591	248 207		
	In all	**1 337 246**	**1 654 196**		
	Cross border trade, net	N/A	- 4 710		
	Total	**1 337 246**	**1 649 487**		
	- Of which bunkered by Danish ships abroad	96 821	254 439		
	- Of which bunkered by Danish planes abroad	2 443	8 015		
Memo:	Total Industries exclusive of bunkering and cross border trade	990 491	1 143 535		

The Danish ships and planes bunkering abroad are part of the industry Transport, storage and communication.

2.3. The air emissions accounts: Data sources and method

The air emissions of the NAMEA type are overall calculated by multiplying emission factors with the direct energy consumption. However, in relation to the assessment of the emissions it is only the use of primary energy (except crude oil and semi-manufactured oil), refined petroleum products and renewable energy (except wind and water power), on which the calculations are based. Consumption of converted types of energy such as electricity and district heat do not cause emissions in itself, i.e. the emissions measured are only caused by the primary energy used to produce these kinds of energy.

Data sources
The primary sources used to establish the emissions accounts are the Danish energy accounts and emission factors and emission inventories obtained from the Danish National Environmental Research Institute (NERI).

Method
Mathematically, the calculations can be expressed by the following: Let E_{ij} be the total amount (in GJ) of energy type i used in industry j or households and let e_{hij} be kilograms of emissions of pollutant h per GJ of energy type i used in sector j. The total emission of pollutant h connected to the use of energy type i in sector j is then EM_{hij} given by:

$$EM_{hij} = E_{ij}e_{hij}$$

h = CO_2, SO_2, NO_x, CO, NH_3, N_2O, CH_4, and NMVOC	(substances)
i = 1, …, 40	(types of energy carrier)
j = 1, …, 130	(industries + households)

E_{ij} is taken directly from the Danish energy accounts, while e_{hij} generally corresponds to the emission factors obtained from NERI. The number of different e_{hij} is limited as the emissions of a single type, h, caused by use of energy type, i, in most cases (but not all), are the same for different industries/households, j. The emission, for example, of CO_2 per unit of gasoline is the same whether gasoline is used in the dairy sector or in the sector for book printing.

Balancing procedure
This calculation gives a theoretic break down of the emissions by industries and households and types of energy products. This theoretic break down of the emissions is afterwards levelled, excepted are the industries *Water transport* and *Air transport*, to the emission totals in the emission inventories submitted by NERI to the United Nations Framework Convention on Climate Change (UNFCCC) and the United Nations Economic Commission for Europe (UNECE) and the European Union (EU).

The reason for the balancing of the emissions to the level in NERI's emission inventories is to make the air emissions accounts fully consistent with NERI's reports to the UNFCCC, UNECE and the EU. Another important argument for the balancing is to account for different abatement technologies in use. This is the case for SO_2 and NO_x from power plants, where emissions to a large extent are measured by monitoring equipment at the power plants (by government regulation and control).

For most industries this emission data correspond to what is reported to the international conventions. Calculation of emissions from renewable energy follows the same procedure as used for non-renewable energy.

Special calculations for road transport
For the use of LPG, motor gasoline and diesel oil the calculation of emissions is done on a more detailed level than described above. The calculation for these three types of energy is based on a breakdown of energy consumption in industries and households into different use categories. That is by types of vehicles.

Integrated in the satellite energy system is a "car system" in which the consumption of gas oil, petrol, and LPG for cars is estimated for each of the 130 industries, household and different types of cars. The car system operates with 189 different types of cars. For each industry/households and type of car the emissions have been estimated, using specific emission factors for the relevant type of car. Calculation of emissions on a detailed level like this gives more reliable estimates for some emission types in cases where emission factors vary from one type of car to another. These estimates are also balanced to the level in NERI's emission inventories.

Emissions and emission coefficients
In order to produce statistics for the NAMEA tables on emissions, it is thus necessary to obtain a set of emission coefficients to be multiplied with the data on energy consumption just described. The coefficients themselves are not published, but it is possible to derive them implicitly from division of the emission matrices by the energy consumption matrices.

The table below shows an aggregated version of the Danish air emissions accounts.

Table 2.2 The Danish air emissions accounts 2002

		CO_2	SO_2	NO_x
		1 000 tonnes		
1	Agriculture, fishing and quarrying	5 369	3	44
2	Manufacturing	8 451	8	16
3	Electricity, gas and water supply	29 144	10	45
4	Construction	1 243	0	18
5	Wholesale and retail trade; hotels, restaurants	1 210	0	6
6	Transport, storage and communication	25 680	440	586
7	Financial intermediation, business activities	437	0	2
8	Public and personal services	912	0	4
	Total industries	**72 446**	**461**	**721**
	Households	12 787	2	38
	Other non-energy related emissions	131	0	0
	In all	**85 364**	**463**	**759**
	Cross border trade, net	- 245	0	- 2
	Total	**85 119**	**463**	**756**
	- Of which energy related emissions caused by industries exclusive of bunkering and cross border trade	49 581	26	174
	- Of which energy related emissions caused by the households	12 787	2	38
	- Of which bunkered by Danish ships abroad	19 846	435	541
	- Of which bunkered by Danish planes abroad	907	0	2
	- Of which emissions from biomass	8 454		
Memo:	Total Industries exclusive of bunkering and cross border trade	51 692	26	177
Memo:	CO_2 – removal, net	- 3 813		

The Danish ships and planes bunkering abroad are part of the industry Transport, storage and communication.

Emissions caused by bunkering
Emissions from fuel oil and jet petroleum bunkered by Danish ships and planes in foreign countries are not included in the calculation of emissions from the energy accounts as described above. Instead, the emissions are calculated by multiplying the fuel use by corresponding emission factors for international sea or air transport. For SO_2 and NO_x, the emission factor for international sea transport is higher than the emission factor for national sea transport.

Emissions from cross border trade
Emissions from cross border trade with motor gasoline and diesel oil are calculated in the same way as emissions from bunkering. The emission factor used to the calculations is approximated to the emission factor for a passenger car.

Non-energy related emissions
The source is NERI's emission inventories. The non-energy related emissions are either connected to the economic activities described in the National accounts or to activities which can not be connected to specific industries or the households.

2.4. Emission data used in the decomposition analysis

However, even though the air emissions accounts account for energy and non-energy related emissions as well as emissions caused by the bunkering abroad and cross border trade it is only the energy-related emissions exclusive of the bunkering abroad and cross border trade which is analysed in the decomposition analysis. The obvious argument for this is that the energy related emissions are connected directly to the economic activity. The argument for not taking the bunkering abroad in to account in the decomposition analysis is that it would be too dominant.

Thus the basis for the decomposition analysis is made up of the items in table 2.3.1. *Of which Energy related emissions caused by industries exclusive of bunkering and cross border trade*, together with *Of which energy related emissions caused by households*.

3. Input-output and National Accounts Data

One of the primary sources of data for a structural decomposition analysis is the input-output (i/o) model. So in the following section the Danish i/o tables and i/o model is described in some detail. Furthermore, a brief overview is given of how the data on energy consumption and emission coefficients that were already described in details in section 2 of this paper is prepared for the analysis.

3.1. Input-output tables and model

The supply side as well as the demand side of the Danish i/o tables are described in detail and linked together in a system of bookkeeping identities, which is fully consistent with the National Accounts. Thus, the input output tables comprise the same 130 industries as the national accounts do at its most disaggregated level. 107 categories of final demand are also included in the input-output tables. In a schematic, fully aggregated form it can be described in the following way

Table 3.1 The structure of the Danish i/o tables

	Intermediate input 1 .. 130	Final demand 1 .. 107	Total
Danish production 1 ... 130	$\mathbf{X^g}$	$\mathbf{F^g}$	**g**
Imports 1 ... 130	$\mathbf{X^m}$	$\mathbf{F^m}$	**m**
Primary factors 1 … 5	**S**	$\mathbf{S^f}$	**s**
Total	**g'**	**f'**	

Note: These matrices are in levels. Both current price tables and tables in fixed 1995 prices are available.

Total output **g** amounts to intermediate goods and services produced in Denmark plus final demand of goods and services produced in Denmark (row sums of $\mathbf{X^g}$ and $\mathbf{F^g}$). Similarly, total imports **m** is distributed between intermediate input and final demand. Primary factors **S** consist mainly of input of labour and capital, but also subsidies and direct and indirect taxes are found here. The column-sums of the primary factors matrix **S** are the gross value added in each of the 130 industries. The matrix $\mathbf{S^f}$ is VAT and other taxes and subsidies. The level of total final demand by category is described by the vector **f'**.

3.2. Import

The import to Denmark is known at the level of 2750 goods and services. They are aggregated to the 130-industry level in the same relative way as the domestically produced goods and services are distributed. If something is imported which is not produced in Denmark, it is assigned an industry code according to its character. A few special categories of import, which cannot be assigned to an existing product or industry, are put in a group of "non-distributable foreign transactions". This import is carried in a 5 by 130 matrix of deliveries to input in production and a 130 by 107 matrix of deliveries to final demand. These additional import matrices are now shown in table 3.1 above in order not to confuse the general picture too much. As the import $\mathbf{X^m}$ and $\mathbf{F^m}$ is classified in the same way as the Danish production and final demand $\mathbf{X^g}$ and $\mathbf{F^g}$ the two sets can be added to get $\mathbf{X} = \mathbf{X^g} + \mathbf{X^m}$ for the intermediate input and $\mathbf{F} = \mathbf{F^g} + \mathbf{F^m}$ for the final demand. If the equivalence between row- and column-sums is to be maintained, the column vector **-m** should be added among the final demand components in **F**.

3.3. Aggregation level

The level of aggregation for production as well as imports is 130. There are 73 categories of private consumption as well as 21 (only 11 before 1993) categories of government consumption and 10 categories of capital formation. Behind these tables, account is being kept of about 2500 goods and services in current and fixed prices. They are used for creating the current as well as the 1995 fixed price i/o tables. The tables have been constructed for the period 1966 to 2000 in both fixed and current prices.

As usual it is assumed that every industry produces only one good, or that the goods they produce all are produced with the same technology. Another thing is that the homogeneity of prices that can be found at the 2500 goods and services level, cannot be maintained at the 130 industry level. It is due to the aggregation process. Therefore the price on delivery from one industry varies between the different uses of it.

3.4. From tables to model

For analytical purposes the i/o table give some valuable information. But it is not always enough. So in order to get a better understanding of detailed structural aspects of the economy it is advantageous to build the tables together in an i/o model where the matrices in the above model are divided by the column sums

Table 3.2 Input output model with endogenous imports

	Intermediate input 1 … 130	Final demand 1 … 107
Danish production 1 ... 130	$\mathbf{A^g}$	$\mathbf{E^g}$
Imports 1 ... 130	$\mathbf{A^m}$	$\mathbf{E^m}$
Primary factors 1 ... 5	$\mathbf{Y}$	$\mathbf{Y^f}$
Total	$\mathbf{i^{g'}}$	$\mathbf{i^{f'}}$

This model is called a model with endogenous import, because it shows all the import transactions explicitly. In a model with exogenous import Danish production and imports will be added together. In the analysis later on we use both types of model in order to be able to compare the results.

In order to do analysis with the model we need to put it on a more usable form. We can write the model as

$$\mathbf{g = A^g g + e^g} \tag{3.1}$$

where the variable names are same as in table 3.2 above, and the only new variable is $\mathbf{e^g}$ which is just a vector of final demand (the row sums of $\mathbf{F^g}$ in table 3.1). Here we can regard $\mathbf{e^g}$ as an exogenous variable and determine the production in each industry in $\mathbf{A^g}$ as the solution to (3.1)

$$\mathbf{g = (I - A^g)^{-1} \cdot e^g} \tag{3.2}$$

where $\mathbf{(I - A^g)^{-1}}$ is the Leontief inverse matrix**.** This equation shows the value of total production in each of the industries in $\mathbf{A^g}$ as a linear function of the supply from the same industries to final demand. Final demand can enter as a matrix instead of a column vector

$$\mathbf{g = (I - A^g)^{-1} \cdot E^g \cdot f} \tag{3.3}$$

where **f** comes from table 3.1 and represents the level of the final demand categories. This model (3.3) can be used to forecast total output, assuming that the technical coefficients are constant. This is not very important in relation to this project, because we are only concerned with the years covered by statistical data. However, because the publication of detailed i/o tables lack behind the publication of more aggregated macroeconomic variables we have in the empirical section of this report forecasted the entire table of coefficients a couple of years to catch up with the newest aggregated data. We shall come back to that in a little while.

The model (3.3) can be subject to a decomposition analysis as it is. In order to explain the changes in industry output in the vector **g,** one would look at the changes in the production structure $(\mathbf{I} - \mathbf{A}^g)^{-1}$, changes in the structure of the final demand $\mathbf{E}^g$ and changes in the level of final demand **f**. But in order to deal with air emissions, this model must be extended with an environmental module.

3.5. Environmental extension of the i/o model

Thus, because the energy matrices are coherent with the national accounts, it is possible to relate energy consumption and emissions with the economic activity at the most detailed industry level. I.e. it is possible to relate the development in physical quantities to the development in the economic activity. The basic definition of an environmental extension is exactly that it relates the basic i/o model with matrices of physical energy consumption and emissions

The simplest environmental extension of the i/o model is through a pre-multiplication of the model (3.3) by a vector of "emission intensity" coefficients. Such coefficients would be obtained by dividing a vector of emissions by industry by the vector of output by industry. Thus, when focus is on CO_2 emissions, the environmentally extended model would be

$$\mathbf{CO_2} = \mathbf{em\_int} \cdot (\mathbf{I} - \mathbf{A}^g)^{-1} \cdot \mathbf{E}^g \cdot \mathbf{f} \qquad (3.4)$$

where, **em_int** = $\mathbf{CO_2}/\mathbf{g}$ is the vector of CO_2 emission intensity coefficients. This model is usable for a decomposition analysis, because it has three different factors that all contribute to the total emissions. We take a closer look at different decomposition models in section 4.4. But first it is necessary to take a closer look at the theory and the methods behind the decomposition analyses.

4. Methodological considerations

There is a number of different techniques that can be used for decomposing the development in emission indicators at the sectoral level. According to Hoekstra and van den Berg (2003) they can be categorized under two general headings; structural decomposition analysis (SDA) and index decomposition analysis (IDA)[3]. Here we shall concentrate on the SDA method.

---

[3] SDA is a generally accepted name for decomposition studies based on input-output models and data. The name IDA is not as generally accepted as SDA, but it is used by Rose and Casler (1996) and adopted in a recent major survey of the area (Ang and Zhang, 2000).

4.1. SDA decomposition analysis – theoretical background

The history of SDA goes back to Leontief and Ford (1972), but traces of it can be found even earlier. Carter (1970) analysed changes in input-output tables over time. Skolka began his work in the second half of the 1970'es leading up to his frequently cited article Skolka (1989). The first well-known Danish contribution is Ploeger (1984). Rose and Casler (1996) carry a more thorough review of the history of SDA.

In Rose and Chen (1991) the SDA method was defined as "the analysis of economic change by means of a set of comparative static changes in key parameters in an input-output table". A number of SDA studies have focused on changes in energy-consumption and changes in emissions of CO_2 and other air-polluting gases. With the purpose of analysing energy demand and emissions, physical data on the environment can be linked to monetary input-output tables either through a product of a number of vectors and matrices representing the "pollution intensity" or through the method of "hybrid units". The latter method allows for the use of monetary as well as physical units in the rows of the input-output tables. This method is considered to be theoretically superior to the intensity factor method if product prices are not uniform across all uses (Hoekstra and van der Berg, 2002). However, it requires some more data-work than the intensity-factor method. Therefore it is more rarely used, but examples can be found in Lin and Polenske (1995), Casler and Rose (1998) and Zheng (2000).

The first SDA studies published, often employed ad-hoc specification of estimating-equations - if equations were presented at all (Rose and Casler, 1996). Since the Rose and Casler (1996) survey, a lot of work has been carried out on the theoretical background of SDA, most of which is presented in Hoekstra and van den Berg (2002).

The idea behind the SDA method can be illustrated as follows in the case of a two-determinant multiplicative function

$$y = x_1 \cdot x_2 \tag{4.1}$$

If we express the left hand side in absolute change terms Δy, we can make two *additive* decompositions, one with the right hand side expressed in absolute terms and one with the right hand side expressed in relative terms. The choice of decomposition form depends on the objective of the analysis. For SDA the adaptive version with both sides expressed in absolute terms is by far the most common. It is also the one that is used in the rest of this paper.

After differentiation of (4.1) by the product rule and using the discrete time approximation, we get the additive decomposition form

$$\Delta y = x_2 \cdot \Delta x_1 + x_1 \cdot \Delta x_2 \tag{4.2}$$

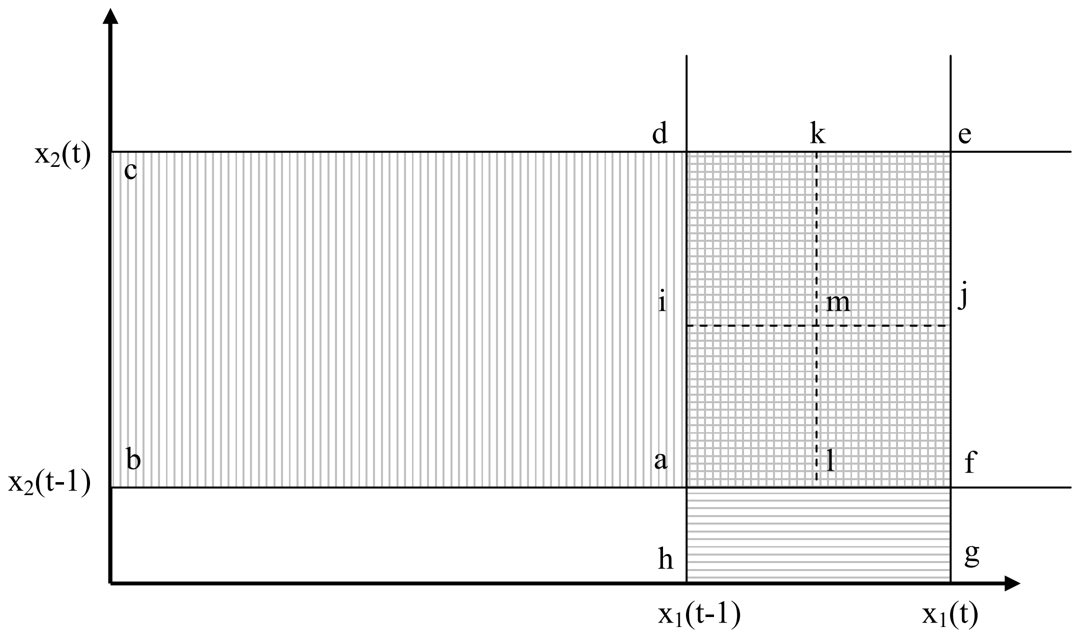
Thus, (4.1) can be decomposed into two parts that depend on the changes in x_1 and x_2. However the choice of weights (here x_2 and x_1) is a very fundamental question, because if we try to rewrite (4.2) in continous time, the identity is not necessarily fulfilled any more. The choice of weights is synonymous with the choice of *index*. We remember that the Laspeyres index means that we use basic year values as weights, while in the Paasche index we use the previous year values as weights. Finally in the Marshall-Edgeworth index we use an average of the two as our weights.

Liu et al. (1992) show that under certain conditions the discrete approximation of a continuous integral function of Δy can be represented by the parametric equation

$$\Delta y \approx (w_1(t-1) + \alpha_1 \cdot \Delta w_1) \cdot \Delta x_1 + (w_2(t-1) + \alpha_2 \cdot \Delta w_2)\Delta x_2 \qquad (4.3)$$

where the w's are weights (that could be x_1 and x_2) and the α's are parameters. The sizes of the weights are determined by their value in period t-1 and t and the parameter α. The choice of the α's determines which index is used. If α is equal to one, only w_1(t) and w_2(t) are used as weights. Thus, we are dealing with the Paasche index. If α is equal to zero, only w_1(t-1) and w_2(t-1) are applied like in the Laspeyres index. Finally if α is equal to 0.5, we are dealing with the Marshall-Edgeworth index; 0.5·w_1(t-1) + 0.5·w_1(t) and 0.5·w_2(t-1) + 0.5·w_2(t). We will try to make this clearer using a graphical presentation inspired by Hoekstra and van den Berg (2002).

Figure 4.1 Additive decomposition of y = x1·x2, discrete time



The total change in y from period t-1 to period t is equal to the total area *bcegha*. Different index methods can be used to calculate the size of this area. Some of them are represented in table 4.1 below.

Table 4.1 Index calculations of Δy

Index	Δy =	Area	Residual
L-L	$\Delta x_1 \cdot x_2(t-1) + \Delta x_2 \cdot x_1(t-1)$	hafg + bcda	$\Delta x_1 \cdot \Delta x_2$ = adef
P-P	$\Delta x_1 \cdot x_2(t) + \Delta x_2 \cdot x_1(t)$	hdeg + bcef	$-\Delta x_1 \cdot \Delta x_2$ = -adef
L-P	$\Delta x_1 \cdot x_2(t-1) + \Delta x_2 \cdot x_1(t)$	hafg + bcef	0
P-L	$\Delta x_1 \cdot x_2(t) + \Delta x_2 \cdot x_1(t-1)$	hdeg + bcda	0
M-E	$\Delta x_1 \cdot 0.5(x_2(t-1)+ x_2(t))$ + $\Delta x_2 \cdot 0.5(x_1(t-1)+ x_1(t))$	hijg + bckl	0

Note: L-L (P-P) refers to a calculation where the Laspeyres (Paasche) index is used for the effects of changes in x_1 as well as in x_2. The L-P and P-L are mixed cases where the effects of changes in x_1 and in x_2 are measured with different indices. M-E refers to the Marshall-Edgeworth index.

From figure 4.1 and table 4.1 it appears that when the Laspeyres index is used for changes in both x_1 and x_2, then the total change in y is underestimated, because the area *adef* is not accounted for. Conversely, when using the Paasche index the area *adef* is counted twice and the total change is overestimated. The reason that Paasche and Laspeyres indices are often used anyway, is that in the case where the relative change is small, i.e. where Δx_1 (Δx_2) is only a small share of x_1 (x_2) the problem is not as crucial as it appears from figure 4.1. However, these cases of over- and underestimation are, what we shall refer to later on as, decompositions with a residual.

The residual term appears first of all in decompositions where the Laspeyres or the Paasche indicies are applied. It represents the so-called mixed effect that arises from a simultaneous change in both components. There are generally two attitudes towards a residual term in the equations

- It is unwanted, so the decomposition must be specified in a way that avoids it, using other indices than pure Laspeyres and Paasche. The residual term is unwanted if we only consider the so-called *isolated effects,* (Seibel, 2003) where changes in each determinant are considered, while all other determinants are being assumed constant.

- It is accepted, and then there are at least three different things to do with it (Seibel, 2003). Firstly, it can simply be neglected, which leads to an incomplete decomposition. This procedure can be justified if the residual is sufficiently small. Secondly, the residual can be distributed among the other determinants. Finally, the residual can be explicitly considered, so that *isolated effects* as well as *mixed effects* are reported.

The emphasis in this report will be on the first type of attitude. So in this situation, an obvious solution is to use different indices to measure the effects of the changes in x_1 and x_2 as it is done in the L-P and P-L cases. The residuals are zero because the area *adef* is counted in both cases, but only once. Thus, the decomposition of y is not unique since there are two different possible decomposition forms. But the result is unique in the two-determinant case. The two decompositions are equivalent and there is no reason why one of them should be preferred to the other.

Another strategy is to apply the Marshall-Edgeworth index, which do not give residuals either. It should be noted, however, that in this case the area *aiml* is counted twice and the area *mkej* is not counted. But in this two-dimensional case they will always be exactly the same size. So the extra *aiml* makes up for the missing *mkej* and they neutralize each other. Thus the decomposition gives no residual in this case.

So if we had only two determinants in our decomposition, we could use either one of the mixed-index cases or the M-E index and get no residuals. However, most SDA studies will incorporate three or more determinants on the right hand side of the equation, and then we have a problem. With e.g. the M-E index it is obvious that what is counted more than once in a SDA with three or more dimensions will only in very special occasions exactly make up for what is not counted. This method is bound to give residuals when the number of determinants is greater than two, because is not complete. The mixed-index method (a mixture of Laspeyres and Paasche indicies) is more promising in the multidimensional case, and we shall return to that later.

The problem is the so-called non-uniqueness, which means that there exist a number of different decomposition forms and that it cannot be decided which one to prefer. It seems that in the literature not very much attention has been paid to this problem. In Dietzenbacher and

Los (1998) it is stated "*… for the economically more meaningful decompositions with a larger number of determinants, the non-uniqueness problem, it's extent and it's implications seem to have been largely neglected"*. In the literature a variety of most often ad-hoc solutions can be found. Lin and Polenske (1995) or Rose and Casler (1998) are mixing Lapeyres and Paasche indices to get rid of residuals. Another example is Wier (1998) and Jakobsen (2000), which, based on Betts (1989) and Fujimagari (1989), take the mean of two decomposition forms, one based on the Lapeyres index and one based the Paasche index. The method gives a residual term.

Another line of authors who has a more thorough and systematic approach to this problem is Dietzenbacher and Los (1998), de Haan (2001) and Seibel (2003). On the basis of these contributions it is possible to derive complete and non-arbitrary decompositions in the n-dimensional case.

4.2. Derivation of estimating equations

The first problem is to write the estimating equations. As the additive decomposition form is used we can apply the so-called additive identity splitting to derive the estimating equations. It involves the addition and subtraction of like terms and rearranging them in the equation. In the case where we have the equation

$$y_t = a_t \cdot b_t \cdot c_t \cdot d_t \tag{4.4}$$

Note that the notation has changed in order to make the equations more readable. Now the subscript t indicates time, and the determinants are differentiated by their name. The additive form is

$$\Delta y = w^a \Delta a + w^b \Delta b + w^c \Delta c + w^d \Delta d$$

Here the *w*'s with superscripts refers to what we might call coefficients or weights. In principle these coefficients could be calculated with econometric methods. But it is also possible to derive them with the structural decomposition method. The additive identity splitting method is used to get an idea of what the *w*'s should be:

$$\begin{aligned}
\Delta y &= y_1 - y_0 \\
\Delta y &= a_1 \cdot b_1 \cdot c_1 \cdot d_1 - a_0 \cdot b_0 \cdot c_0 \cdot d_0 \\
&= \Delta a \cdot b_1 \cdot c_1 \cdot d_1 + a_0 \cdot b_1 \cdot c_1 \cdot d_1 - a_0 \cdot b_0 \cdot c_0 \cdot d_0 \\
&= \Delta a \cdot b_1 \cdot c_1 \cdot d_1 + a_0 \cdot \Delta b \cdot c_1 \cdot d_1 + a_0 \cdot b_0 \cdot c_1 \cdot d_1 - a_0 \cdot b_0 \cdot c_0 \cdot d_0 \\
&= \Delta a \cdot b_1 \cdot c_1 \cdot d_1 + a_0 \cdot \Delta b \cdot c_1 \cdot d_1 + a_0 \cdot b_0 \cdot \Delta c \cdot d_1 + a_0 \cdot b_0 \cdot c_0 \cdot d_1 - a_0 \cdot b_0 \cdot c_0 \cdot d_0 \\
&= \Delta a \cdot b_1 \cdot c_1 \cdot d_1 + a_0 \cdot \Delta b \cdot c_1 \cdot d_1 + a_0 \cdot b_0 \cdot \Delta c \cdot d_0 + a_0 \cdot b_0 \cdot c_0 \cdot \Delta d
\end{aligned} \tag{4.5}$$

So now we have a decomposition form, with four terms. Each of the four terms expresses the contribution of the Δ-component to the total change in *y*. In the first term the coefficient attached to Δa is $b_1 \cdot c_1 \cdot d_1$, for Δb it is $a_0 \cdot c_1 \cdot d_1$, and so on for Δc and Δd. We notice a pattern, where the Δ runs from left to right and all coefficients to the right of the Δ-component are counted in the target-year value and all the coefficients to the left of the Δ-component are counted in basic year values. This decomposition form is *complete*, meaning that it has no residual. However, this form is not unique. It is just one of many decompositions. The derivation of the decomposition equation above arbitrarily assumed that the order of the determinants was *abcd*, but it could just as well have been *cadb*. If we follow the principles of (4.5) we will have Δc in the first term and Δa in the next and so on. Dietzenbacher and Los (1998) show that in the general *n*-determinants case there is *n!* different forms[4]. In this case we would have 4!=24 different forms.

Figure 4.2 All 24 decompositions of y=abcd

$\Delta y = \Delta a * b_1 * c_1 * d_1 + a_0 * \Delta b * c_1 * d_1 + a_0 * b_0 * \Delta c * d_1 + a_0 * b_0 * c_0 * \Delta d$	a b c d
$\Delta y = \Delta a * b_1 * d_1 * c_1 + a_0 * \Delta b * d_1 * c_1 + a_0 * b_0 * \Delta d * c_1 + a_0 * b_0 * d_0 * \Delta c$	a b d c
$\Delta y = \Delta a * c_1 * b_1 * d_1 + a_0 * \Delta c * b_1 * d_1 + a_0 * c_0 * \Delta b * d_1 + a_0 * c_0 * b_0 * \Delta d$	a c b d
$\Delta y = \Delta a * c_1 * d_1 * b_1 + a_0 * \Delta c * d_1 * b_1 + a_0 * c_0 * \Delta d * b_1 + a_0 * c_0 * d_0 * \Delta b$	a c d b
$\Delta y = \Delta a * d_1 * b_1 * c_1 + a_0 * \Delta d * b_1 * c_1 + a_0 * d_0 * \Delta b * c_1 + a_0 * d_0 * b_0 * \Delta c$	a d b c
$\Delta y = \Delta a * d_1 * c_1 * b_1 + a_0 * \Delta d * c_1 * b_1 + a_0 * d_0 * \Delta c * b_1 + a_0 * d_0 * c_0 * \Delta b$	a d c b
$\Delta y = \Delta b * a_1 * c_1 * d_1 + b_0 * \Delta a * c_1 * d_1 + b_0 * a_0 * \Delta c * d_1 + b_0 * a_0 * c_0 * \Delta d$	b a c d
$\Delta y = \Delta b * a_1 * d_1 * c_1 + b_0 * \Delta a * d_1 * c_1 + b_0 * a_0 * \Delta d * c_1 + b_0 * a_0 * d_0 * \Delta c$	b a d c
$\Delta y = \Delta b * c_1 * a_1 * d_1 + b_0 * \Delta c * a_1 * d_1 + b_0 * c_0 * \Delta a * d_1 + b_0 * c_0 * a_0 * \Delta d$	b c a d
$\Delta y = \Delta b * c_1 * d_1 * a_1 + b_0 * \Delta c * d_1 * a_1 + b_0 * c_0 * \Delta d * a_1 + b_0 * c_0 * d_0 * \Delta a$	b c d a
$\Delta y = \Delta b * d_1 * a_1 * c_1 + b_0 * \Delta d * a_1 * c_1 + b_0 * d_0 * \Delta a * c_1 + b_0 * d_0 * a_0 * \Delta c$	b d a c
$\Delta y = \Delta b * d_1 * c_1 * a_1 + b_0 * \Delta d * c_1 * a_1 + b_0 * d_0 * \Delta c * a_1 + b_0 * d_0 * c_0 * \Delta a$	b d c a
$\Delta y = \Delta c * a_1 * b_1 * d_1 + c_0 * \Delta a * b_1 * d_1 + c_0 * a_0 * \Delta b * d_1 + c_0 * a_0 * b_0 * \Delta d$	c a b d
$\Delta y = \Delta c * a_1 * d_1 * b_1 + c_0 * \Delta a * d_1 * b_1 + c_0 * a_0 * \Delta d * b_1 + c_0 * a_0 * d_0 * \Delta b$	c a d b
$\Delta y = \Delta c * b_1 * a_1 * d_1 + c_0 * \Delta b * a_1 * d_1 + c_0 * b_0 * \Delta a * d_1 + c_0 * b_0 * a_0 * \Delta d$	c b a d
$\Delta y = \Delta c * b_1 * d_1 * a_1 + c_0 * \Delta b * d_1 * a_1 + c_0 * b_0 * \Delta d * a_1 + c_0 * b_0 * d_0 * \Delta a$	c b d a
$\Delta y = \Delta c * d_1 * a_1 * b_1 + c_0 * \Delta d * a_1 * b_1 + c_0 * d_0 * \Delta a * b_1 + c_0 * d_0 * a_0 * \Delta b$	c d a b
$\Delta y = \Delta c * d_1 * b_1 * a_1 + c_0 * \Delta d * b_1 * a_1 + c_0 * d_0 * \Delta b * a_1 + c_0 * d_0 * b_0 * \Delta a$	c d b a
$\Delta y = \Delta d * a_1 * b_1 * c_1 + d_0 * \Delta a * b_1 * c_1 + d_0 * a_0 * \Delta b * c_1 + d_0 * a_0 * b_0 * \Delta c$	d a b c
$\Delta y = \Delta d * a_1 * c_1 * b_1 + d_0 * \Delta a * c_1 * b_1 + d_0 * a_0 * \Delta c * b_1 + d_0 * a_0 * c_0 * \Delta b$	d a c b
$\Delta y = \Delta d * b_1 * a_1 * c_1 + d_0 * \Delta b * a_1 * c_1 + d_0 * b_0 * \Delta a * c_1 + d_0 * b_0 * a_0 * \Delta c$	d b a c
$\Delta y = \Delta d * b_1 * c_1 * a_1 + d_0 * \Delta b * c_1 * a_1 + d_0 * b_0 * \Delta c * a_1 + d_0 * b_0 * c_0 * \Delta a$	d b c a
$\Delta y = \Delta d * c_1 * a_1 * b_1 + d_0 * \Delta c * a_1 * b_1 + d_0 * c_0 * \Delta a * b_1 + d_0 * c_0 * a_0 * \Delta b$	d c a b
$\Delta y = \Delta d * c_1 * b_1 * a_1 + d_0 * \Delta c * b_1 * a_1 + d_0 * c_0 * \Delta b * a_1 + d_0 * c_0 * b_0 * \Delta a$	d c b a

The rightmost column of figure 4.2 shows the 24 permutations of the determinants *a,b,c* and *d* that has been used to generate the 24 equations. All of these n! decomposition equations yield exactly the same value of Δy. Different coefficients are attached to the n components, but the derivation in (4.5) ensures identical values of Δy and no residuals.

But the size of the contribution from Δa, Δb, Δc and Δd differ across the equations. The difference in coefficients mean that dependent on which of the n! decompositions we look at, we can see quite different contributions from the same determinant to the change in *y*. As shown in de Haan (2001) (and also later in this report) there can be a huge difference between any of the 24 suggestions to what the contribution of one determinant might be and the mean of all the 24 suggestions. In other words, the variance can be very large. De Haan reports a variance of –60% to +70% with respect to the mean. That is an evidence of how wrong it would be to arbitrarily pick just one of the n! equations and calculate the contribution of the n factors to the change in *y*.

Dietzenbacher and Los (1998) suggest that a way to reduce the variance is to look at the mean of so-called "mirror images". Mirror images are pairs of permutations where the time period indication on the coefficients attached to each difference term are exactly opposite, like e.g. in line 1 and line 24 in figure 4.2 above. The n equations comprise n!/2 such pairs. For the two components in such a pair, the deviation from the mean goes in opposite directions. Thus, the mean of the two components in a pair are quite close to the overall mean. Actually, de Haan

[4] As noted by de Haan (2001), it has been showed that there exists a lot more than (*n!*) decomposition form. Thus all equations with more than one difference term Δ could be considered too. However, the economic interpretation of those terms is not straightforward.

(2001) reports deviations of only 0-1% with very few exceptions, as opposed to the –60% to +70% mentioned earlier.

Dietzenbacher and Los (1998) suggested an improvement to the polar-case solution. It was to calculate the mean of all the *n!* decomposition forms in figure 4.2. The results of their example show that there is substantial variation in the outcome of their 24 decomposition forms, just like it was found in de Haan (2001). Their advice is therefore to calculate all *n!* forms and to publish standard deviations together with the means. In order to calculate the mean of the four determinants in the example in figure 4.2 above, the equations need to be sorted to get the *Δ*'s of the same determinants put into the same column.

Figure 4.3 All 24 decompositions of y=abcd, sorted

$\Delta y = \Delta a*b_1*c_1*d_1 + a_0*\Delta b*c_1*d_1 + a_0*b_0*\Delta c*d_1 + a_0*b_0*c_0*\Delta d$	a b c d
$\Delta y = \Delta a*b_1*c_1*d_1 + a_0*\Delta b*c_1*d_1 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_1*\Delta d$	a b d c
$\Delta y = \Delta a*b_1*c_1*d_1 + a_0*\Delta b*c_0*d_1 + a_0*b_1*\Delta c*d_1 + a_0*b_0*c_0*\Delta d$	a c b d
$\Delta y = \Delta a*b_1*c_1*d_1 + a_0*\Delta b*c_0*d_0 + a_0*b_1*\Delta c*d_1 + a_0*b_1*c_0*\Delta d$	a c d b
$\Delta y = \Delta a*b_1*c_1*d_1 + a_0*\Delta b*c_0*d_1 + a_0*b_0*\Delta c*d_0 + a_0*b_1*c_1*\Delta d$	a d b c
$\Delta y = \Delta a*b_1*c_1*d_1 + a_0*\Delta b*c_0*d_0 + a_0*b_1*\Delta c*d_0 + a_0*b_1*c_1*\Delta d$	a d c b
$\Delta y = \Delta a*b_0*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_0*b_0*\Delta c*d_1 + a_0*b_0*c_0*\Delta d$	b a c d
$\Delta y = \Delta a*b_0*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_1*\Delta d$	b a d c
$\Delta y = \Delta a*b_0*c_0*d_1 + a_1*\Delta b*c_1*d_1 + a_1*b_0*\Delta c*d_1 + a_0*b_0*c_0*\Delta d$	b c a d
$\Delta y = \Delta a*b_0*c_0*d_0 + a_1*\Delta b*c_1*d_1 + a_1*b_0*\Delta c*d_1 + a_1*b_0*c_0*\Delta d$	b c d a
$\Delta y = \Delta a*b_0*c_1*d_0 + a_1*\Delta b*c_1*d_1 + a_0*b_0*\Delta c*d_0 + a_1*b_0*c_1*\Delta d$	b d a c
$\Delta y = \Delta a*b_0*c_0*d_0 + a_1*\Delta b*c_1*d_1 + a_1*b_0*\Delta c*d_0 + a_1*b_0*c_1*\Delta d$	b d c a
$\Delta y = \Delta a*b_1*c_0*d_1 + a_0*\Delta b*c_0*d_1 + a_1*b_1*\Delta c*d_1 + a_0*b_0*c_0*\Delta d$	c a b d
$\Delta y = \Delta a*b_1*c_0*d_1 + a_0*\Delta b*c_0*d_0 + a_1*b_1*\Delta c*d_1 + a_0*b_1*c_0*\Delta d$	c a d b
$\Delta y = \Delta a*b_0*c_0*d_1 + a_1*\Delta b*c_0*d_1 + a_1*b_1*\Delta c*d_1 + a_0*b_0*c_0*\Delta d$	c b a d
$\Delta y = \Delta a*b_0*c_0*d_0 + a_1*\Delta b*c_0*d_1 + a_1*b_1*\Delta c*d_1 + a_1*b_0*c_0*\Delta d$	c b d a
$\Delta y = \Delta a*b_1*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_0*\Delta d$	c d a b
$\Delta y = \Delta a*b_0*c_0*d_0 + a_1*\Delta b*c_0*d_0 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_0*\Delta d$	c d b a
$\Delta y = \Delta a*b_1*c_1*d_0 + a_0*\Delta b*c_1*d_0 + a_0*b_0*\Delta c*d_0 + a_1*b_1*c_1*\Delta d$	d a b c
$\Delta y = \Delta a*b_1*c_1*d_0 + a_0*\Delta b*c_0*d_0 + a_0*b_1*\Delta c*d_0 + a_1*b_1*c_1*\Delta d$	d a c b
$\Delta y = \Delta a*b_0*c_1*d_0 + a_1*\Delta b*c_1*d_0 + a_0*b_0*\Delta c*d_0 + a_1*b_1*c_1*\Delta d$	d b a c
$\Delta y = \Delta a*b_0*c_0*d_0 + a_1*\Delta b*c_1*d_0 + a_1*b_0*\Delta c*d_0 + a_1*b_1*c_1*\Delta d$	d b c a
$\Delta y = \Delta a*b_1*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_1*b_1*\Delta c*d_0 + a_1*b_1*c_1*\Delta d$	d c a b
$\Delta y = \Delta a*b_0*c_0*d_0 + a_1*\Delta b*c_0*d_0 + a_1*b_1*\Delta c*d_0 + a_1*b_1*c_1*\Delta d$	d c b a

It is permitted to rearrange the four products in each line, just as it is permitted to rearrange the terms inside each of the four products. That will be necessary if we work with vectors and matrices that have to come in a certain order because of their different dimensions.

Now all the *Δ*'s are in the column where they belong. Software code can be generated to calculate these equations and to take means and standard deviations on each of the columns. However, if we want 6,7 or 8 determinants in the equations, the number of equations will rise to 720, 5040 and 40320 respectively, which even for a modern computer can be quite time-consuming to compute.

But, actually it is not necessary to do so. In order to reduce the number, the first step is to accept the fact that we actually do not need to calculate all the *n Δy*'s once we have convinced ourselves that they are all equal. We only need the means of the *n* columns. That allows us to sort the equations in the vertical direction as well in order to get the result shown I figure 4.4.

Figure 4.4. All 24 decompositions of y=abcd, sorted vertically and horizontally

$$
\begin{aligned}
&\Delta y = 1/24 * [\\
&\Delta a*b_0*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_0*\Delta d +\\
&\Delta a*b_0*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_0*\Delta d +\\
&\Delta a*b_0*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_0*\Delta d +\\
&\Delta a*b_0*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_0*\Delta d +\\
&\Delta a*b_0*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_0*\Delta d +\\
&\Delta a*b_0*c_0*d_0 + a_0*\Delta b*c_0*d_0 + a_0*b_0*\Delta c*d_0 + a_0*b_0*c_0*\Delta d +\\
&\Delta a*b_0*c_0*d_1 + a_0*\Delta b*c_0*d_1 + a_0*b_0*\Delta c*d_1 + a_0*b_0*c_1*\Delta d +\\
&\Delta a*b_0*c_0*d_1 + a_0*\Delta b*c_0*d_1 + a_0*b_0*\Delta c*d_1 + a_0*b_0*c_1*\Delta d +\\
&\Delta a*b_0*c_1*d_0 + a_0*\Delta b*c_0*d_1 + a_0*b_1*\Delta c*d_0 + a_0*b_1*c_0*\Delta d +\\
&\Delta a*b_0*c_1*d_0 + a_0*\Delta b*c_1*d_0 + a_0*b_1*\Delta c*d_0 + a_0*b_1*c_0*\Delta d +\\
&\Delta a*b_0*c_1*d_1 + a_0*\Delta b*c_1*d_1 + a_0*b_1*\Delta c*d_1 + a_0*b_1*c_1*\Delta d +\\
&\Delta a*b_0*c_1*d_1 + a_0*\Delta b*c_1*d_1 + a_0*b_1*\Delta c*d_1 + a_0*b_1*c_1*\Delta d +\\
&\Delta a*b_1*c_0*d_0 + a_1*\Delta b*c_0*d_0 + a_1*b_0*\Delta c*d_0 + a_1*b_0*c_0*\Delta d +\\
&\Delta a*b_1*c_0*d_0 + a_1*\Delta b*c_0*d_0 + a_1*b_0*\Delta c*d_0 + a_1*b_0*c_0*\Delta d +\\
&\Delta a*b_1*c_0*d_1 + a_1*\Delta b*c_0*d_1 + a_1*b_0*\Delta c*d_1 + a_1*b_0*c_1*\Delta d +\\
&\Delta a*b_1*c_0*d_1 + a_1*\Delta b*c_0*d_1 + a_1*b_0*\Delta c*d_1 + a_1*b_0*c_1*\Delta d +\\
&\Delta a*b_1*c_1*d_0 + a_1*\Delta b*c_1*d_0 + a_1*b_1*\Delta c*d_0 + a_1*b_1*c_0*\Delta d +\\
&\Delta a*b_1*c_1*d_0 + a_1*\Delta b*c_1*d_0 + a_1*b_1*\Delta c*d_0 + a_1*b_1*c_0*\Delta d +\\
&\Delta a*b_1*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_1*\Delta d +\\
&\Delta a*b_1*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_1*\Delta d +\\
&\Delta a*b_1*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_1*\Delta d +\\
&\Delta a*b_1*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_1*\Delta d +\\
&\Delta a*b_1*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_1*\Delta d +\\
&\Delta a*b_1*c_1*d_1 + a_1*\Delta b*c_1*d_1 + a_1*b_1*\Delta c*d_1 + a_1*b_1*c_1*\Delta d\ ]
\end{aligned}
$$

Notice that it is no longer possible to calculate Δy in each line. However, the total sum of the 24*4=96 terms are still equal to 24*Δy as it can be seen in figure 4.4. Notice the same pattern in all 4 columns. In the first column in figure 4.4 we see that many of the coefficients attached to the *Δa*'s are equal. The coefficient $b_0*c_0*d_0$ appears 6 times just like the coefficient $b_1*c_1*d_1$ does. Between the 6 identical in the top and the bottom there are 6 pairs of equals.

In each column there are n-1 different coefficients. They can be from two different periods (0 for basic year and 1 for target year). That leaves us with $2^{(n-1)}$ different coefficients to attach to the *Δ*–component ($2^{(4-1)} = 8$ in this example). Thus, in our example, we have only 8 different coefficients in a column of 24 equations. If we take advantage of these findings, we can cut down on the number of equations and increase the speed of computation dramatically. The same types of findings are mentioned in de Haan (2001), but it has been formalized to apply for any *n* dimensional decomposition in Seibel (2003). The following builds on Seibel (2003).

We know now that we must calculate the $2^{(n-1)}$ coefficients. But, what we do not know is what weight should be attached to each of them before we calculate the mean. We can use some mathematics to sort that out. The weight is dependent on two things. Besides the number of determinants n, it also depends on the distribution between base year values (subscript 0) and target year values (subscript 1) in the coefficient. We learned from figure 4.4 that the two coefficients that consist of either only subscript 0 values or subscript 1 values are represented six times. Actually, if we take a closer look, we can see that also coefficients that consist of one subscript 0 value and two subscript 1 values are represented six times. They are just in three different forms $b_0*c_0*d_1$ and $b_0*c_1*d_0$ and $b_1*c_0*d_0$.

Now, we let *k* represent the number of subscript 0 values in a coefficient. So *k* runs from 0 to *n-1*. Then, conversely, *n-1-k* is the number of subscript 1 values in the same coefficient.

Firstly, we would like to know how many different coefficients there is for each value of k. Statistical theory gives the answer. For each *k* the number of coefficients is

$$\frac{(n-1)!}{[(n-1-k)! \cdot k!]} \tag{4.6}$$

In our example from above this gives us the results shown in table 4.2 below. We see that when we let *k* run from 0 to 3 the total number of coefficients 1+3+3+1 = 8 equals $2^{(4-1)}$. Thus, there is one way to write the coefficient when *k* equals zero, 3 different ways when *k* equals one, and when it equals 2. Finally, there is only one way to write it, when *k* equals three.

The next step is to find out how many times each of these coefficients appear as weight for the *Δ*-term in the *n!* equations. Our *k* can have *n* different values (runs from 0 to n-1), so there must be *n* different types of coefficients, when type is determined by the size of *k*. They all have to be represented an equal number of times among the n! equations. That means that each value of *k* must be represented (n! / n) = (n-1)! times among the n! equations. This number must for every value of k be divided between the numbers of different coefficients based on this particular k as calculated by (4.6)

$$\frac{(n-1)!}{\frac{(n-1)!}{[(n-1-k)! \cdot k!]}} = (n-1-k)! \cdot k! \tag{4.7}$$

In our example the result of (4.6) and (4.7) with n=4 gives the following table

Table 4.2. Number of different coefficients and their weights, when n=4

k	Number of different coefficients for n=4, given k *(n-1)! / [(n-1-k)!·k!]*	Weight *(n-1-k)!·k!*
0	1	6
1	3	2
2	3	2
3	1	6

By multiplying columns 2 and 3 we see that each value of *k* will be represented 6 times in this example. With this knowledge we can create a matrix of subscripts 0 and 1 for the $2^{(n-1)}$ different coefficients. For our n=4 example it would look like

Table 4.3 Subscripts for the 3 components in the $2^{(4-1)}$ coefficients

k	Subscripts for the components in the coefficient			Weight
	first	second	third	
0	0	0	0	6
1	0	0	1	2
	0	1	0	2
	1	0	0	2
2	0	1	1	2
	1	1	0	2
	1	0	1	2
3	1	1	1	6

Now we can use the matrix of subscripts marked as the slightly shaded area in table 4.3 to write a new set of equations to replace the ones in figure 4.4 above

Figure 4.5 Δy calculated as the average of all 24 decompositions represented by 8 different decompositions of y, with appropriate weights

$$\begin{aligned}
\Delta y = \ & 1/24 * [\\
& \{6 * \Delta a*b_0*c_0*d_0 + 6 * a_0*\Delta b*c_0*d_0 + 6 * a_0*b_0*\Delta c*d_0 + 6 * a_0*b_0*c_0*\Delta d\} + \\
& \{2 * \Delta a*b_0*c_0*d_1 + 2 * a_0*\Delta b*c_0*d_1 + 2 * a_0*b_0*\Delta c*d_1 + 2 * a_0*b_0*c_1*\Delta d\} + \\
& \{2 * \Delta a*b_0*c_1*d_0 + 2 * a_0*\Delta b*c_0*d_1 + 2 * a_0*b_1*\Delta c*d_0 + 2 * a_0*b_1*c_0*\Delta d\} + \\
& \{2 * \Delta a*b_1*c_0*d_0 + 2 * a_1*\Delta b*c_0*d_0 + 2 * a_1*b_0*\Delta c*d_0 + 2 * a_1*b_0*c_0*\Delta d\} + \\
& \{2 * \Delta a*b_0*c_1*d_1 + 2 * a_0*\Delta b*c_1*d_1 + 2 * a_0*b_1*\Delta c*d_1 + 2 * a_0*b_1*c_1*\Delta d\} + \\
& \{2 * \Delta a*b_1*c_1*d_0 + 2 * a_1*\Delta b*c_1*d_0 + 2 * a_1*b_1*\Delta c*d_0 + 2 * a_1*b_1*c_0*\Delta d\} + \\
& \{2 * \Delta a*b_1*c_0*d_1 + 2 * a_1*\Delta b*c_0*d_1 + 2 * a_1*b_0*\Delta c*d_1 + 2 * a_1*b_0*c_1*\Delta d\} + \\
& \{6 * \Delta a*b_1*c_1*d_1 + 6 * a_1*\Delta b*c_1*d_1 + 6 * a_1*b_1*\Delta c*d_1 + 6 * a_1*b_1*c_1*\Delta d\} \ ]
\end{aligned}$$

Now the final step is to calculate the size of the total contribution from each of the four determinants to the total change in *y* as the average of all 24 decompositions represented by the 8 different decompositions of *y* presented in figure 4.5. To do this, we must look at the columns isolated from each other. So the total change in *y* is equal to

$$\Delta y = w^1 \cdot \Delta a + w^2 \cdot \Delta b + w^3 \cdot \Delta c + w^4 \cdot \Delta d$$

where

$$\begin{aligned}
w^1 \cdot \Delta a = 1/24 * \{ & (6 * \Delta a*b_0*c_0*d_0) + \\
& (2 * \Delta a*b_0*c_0*d_1) + \\
& (2 * \Delta a*b_0*c_1*d_0) + \\
& (2 * \Delta a*b_1*c_0*d_0) + \\
& (2 * \Delta a*b_0*c_1*d_1) + \\
& (2 * \Delta a*b_1*c_1*d_0) + \\
& (2 * \Delta a*b_1*c_0*d_1) + \\
& (6 * \Delta a*b_1*c_1*d_1) \quad \}
\end{aligned}$$

and similarly for the other three columns.

It is now reasonably easy to use the framework outlined above to make a decomposition of models with 2 to n determinants. The number of necessary decomposition equations is reduced from $n!$ to $2^{(n-1)}$, which for large systems reduces the computational requirements dramatically, while exactly preserving the results.

4.3. Conclusion

Structural decomposition analysis has undergone a considerable development in the literature in recent years. Theory and methods have been developed in the direction of complete decompositions with no residuals and more accurate estimates of contributions from the determinants. There are, however, some objections to this method, one of which is the question of *dependent determinants.* It is investigated in Dietzenbacher and Los (2000) how dependency between the determinants, which is actually very common, may affect the results of a SDA. It is indicated that dependencies may cause a bias in the results in certain SDA studies. A new decomposition method that does not suffer from these drawbacks are presented, and in a case study of the Dutch economy it is showed that results obtained with the new method may differ substantially from results obtained with the more traditional method. For future work it would be valuable to take a closer look at this new method. However, the method outlined in the pages above, will in spite of this new development be used in the empirical part of this paper.

5. Setting up Danish Decomposition Analyses

With the necessary data at hand and the decomposition methodology carefully reviewed, it is now possible to derive some models that we can use for the empirical analysis.

The principle behind the derivation is that we have our basic i/o model (3.3) and we premultiply it with a block of energy end environmental matrices and vectors. The overall property of this block must be the same as for the vector **em_int** in equation (3.4) above. When pre-multiplied on the model (3.3) we should be left with a scalar or vector of emissions. Therefore, it is required that the block in general has emissions as the numerator and total output as the denominator. Another requirement is that it must have the same row-dimension as the inverted matrix $\mathbf{(I - A^g)^{-1}}$ so they fit together in a matrix multiplication.

If we have that **A** is the outcome of the basic i/o model i.e. normally total output, and the vector **b** is emissions we have that

$$b = \left(\frac{b}{A}\right)A \tag{5.1}$$

This can be decomposed into more vectors and matrices to the left as long as the basic property of (5.1) still holds

$$b = \left(\frac{b}{c}\right)\left(\frac{c}{d}\right)\left(\frac{d}{e}\right)\left(\frac{e}{A}\right)A \tag{5.2}$$

We see that the **c**, **d** and **e** variables offset each other in the final result. However they can still be valuable determinants in a decomposition analysis and shed some light upon the reasons for the observed changes over time in **b**.

As we have available consistent time series of all data from 1980 to 2002, decomposition has been carried out of changes in emissions between 1980 and every single year from 1981 to 2002 subsequently. So while the base year is kept constant at 1980, the target year gradually runs through the entire time span. Doing it this way we get consistent annual time series of results.

5.1. Basic model

The basic model that are used for decomposition in this report is the following

$$\underset{(130\times1)}{\mathbf{emis}} = \underset{(130\times40)}{\mathbf{emcoef}} \,\#\, \underset{(130\times40)}{\mathbf{emix}} \,\#\, \underset{(130\times1)}{\mathbf{enint}} \cdot \underset{(40\times1)}{\mathbf{summa}} \,\#\, \underset{(130\times130)}{\mathbf{(I - A^g)^{-1}}} \cdot \underset{(130\times97)}{\mathbf{FDstruct}} \cdot \underset{(97\times1)}{\mathbf{FDlevel}} \tag{5.3}$$

Here the symbol # indicates element-by-element matrix multiplication and the symbol · indicates ordinary matrix multiplication. Obviously, the same model can be used for analyses of CO_2, SO_2 or NO_x emissions respectively. It just requires that the **emis** and the **emcoef** variables be updated with information on the pollutant in question. The elements in the equation (5.3) is the following

emis is a 130×1 vector of total emission of CO_2, SO_2 and NO_x by industry.

emcoef is a 130×40 matrix of emission coefficients for CO_2, SO_2 and NO_x. There is a coefficient for the emission from each of the 130 industries of each of the 40 energy carriers. Dividing the 130×40 emission matrices from the NAMEA system by the 130×40 energy consumption matrices in gigajoule creates the **emcoef** matrix. The data for this variable is readily available from 1980 through 2001, which holds true for the **emix** and **enint** variables as well.

emix is also a 130×40 matrix. It is created by an element-by-element division of the matrix of energy consumption by its own row sums. Therefore the row sums of the **emix** matrix are one. It explains the weight with which the 40 energy carriers are used by each of the 130 industries. Thus, this matrix registers changes in the input of energy towards more or less polluting energy carriers.

enint is a 130×1 vector of energy intensities by industry. A division of the row sums of the energy consumption matrix by the total output makes this vector. A change in this variable over time indicates to which extent the various industries have been able to change the production processes in the direction of more efficient use of it energy input.

summa is a 40×1 summation vector, which is necessary to insert, because after multiplication of the first three components the dimension of the matrix is 130×40, which must be 1×130 in order to be compatible with the $\mathbf{(I\text{-}A)^{-1}}$ matrix, which is the next in line. It is possible to avoid this summation vector if e.g. the **emcoef** vector is aggregated to 1×40 and the next matrices are transposed. However, we have found that there are some differences in the results if such a vector is used instead, because aggregation causes loss of detail. In the results of the analysis there is no effect from this vector what so ever, because it does not change at all over time. Thus it is just a tool in the analysis, not a real determinant.

$\mathbf{(I\text{-}\ A^g)^{-1}}$ is the 130×130 inverted matrix of intermediate deliveries. Notice the superscript g as in (3). It indicates that we are dealing with domestically produced intermediate input. In some analyses the focus is on global emissions generated by Danish final demand, in which case the matrix is **A,** indicating that domestic and import matrices are added together. Such addition is then done for the matrices of domestic and imported final demand as well. Please refer to tables 3.1 and 3.2 to see that when we multiply the summed coefficient matrices with the level of final demand we end up with something more than total output **g**, namely **g+m**. Normally what is done to circumvent this problem, is to put an extra column vector in the final demand matrix, with the value **–m**. That will secure the original value of the row sums to be **g**. However in this situation it does not really matter, so we just keep the total **g+m,** because it will then gives us the global emissions from the Danish final demand. This is only possible, because we have made the courageous assumption that final demand goods and services produced abroad are produced with exactly the same energy consumption, emission coefficients and thus emissions, as if it had been produced in Denmark. Naturally, this assumption does not hold, but it is quite difficult to do anything else. But with more and more NAMEA tables appearing in the EU countries it should be possible to collect a sufficient amount of data to improve this part of the model. Actually, in section 5 of this report we have a section that looks at emissions generated by electricity imported from other Nordic countries. Unfortunately it has not been possible within the frames of this project to introduce this new information into the right matrices and to see the consequences of it.

In terms of availability of data for this variable it is limited by the time-lag in publication of i/o tables. At this point in time the latest version of i/o tables is 2000. But, as all the energy and environmental data in this project is published through 2002, a lot of effort has been put into a forecast of i/o tables for 2001 and 2002 as well, in order to facilitate decomposition analyses through 2002. The traditional and easiest way to forecast i/o coefficient matrices is to use the latest available published tables and forecast them as constants. Initially, such method was used in this project, but combined with row and column sums in terms of published statistics on macroeconomic aggregates, it would give a quite incredible development in many cells in the 2000 and 2001 i/o tables. Therefore it was decided to do a full "rAs" balancing of those two years[5]. When the balanced coefficient matrices were used in the decomposition analyses, the results had no obvious data-breaks and seemed more sensible and credible than with the constant forecast.

FDstruct is a 130×97 matrix of final demand coefficients. As mentioned previously there are 73 groups of private consumption and then also investment, government consumption and export. Only from 1993 and onwards we have the 107 groups of final demand as indicated by tables 3.1 and 3.2, so those years have been aggregated to comply with the earlier years. Changes in the preferences of the final consumers will be represented by changes in this matrix.

FDlevel is a 97×1 matrix of the level of final demand. The general growth of the economy is quite well represented by this vector. In some studies this vector is converted into shares that sum to one, of the total final demand and then a scalar of the total final demand is added as the last determinant of the decomposition.

With this equation it is possible to use the decomposition method laid out in the previous section to get some interesting results. This decomposition results in a 130H1 vector. It can either be summed over all industries to tell a story about changes in total emissions from industries or groups of it can be summed to tell a story about different sectors of the economy like agriculture, industry, transport and so on. Both types of results can be seen in section 6 below.

[5] The rAs technique is a so-called biproportional adjustment method for updating and interpretating change in input-output accounts. The method generates a new input output coefficients table using a prior year table in conjunction with information on the current year row and column sums. It is a mathematical optimisation algorithm, and, as Bacharach (1970) noted, "*one estimates the unknown matrix as that value which, if realized, would occasion the least 'surprise' in view of the prior*". For a recent overview of methods for updating input-output matrices, see Jackson and Murray (2002).

5.2. A final demand variant

Because we are also interested to see what resulted in the observed changes in emissions caused by different groups of final demand we have made a slightly different model

$$\underset{(1\times97)}{\mathbf{emis}} = \underset{(1\times40)}{\mathbf{emcoef2}} \cdot \underset{(130\times40)'}{\mathbf{emix'}} \# \underset{(130\times1)'}{\mathbf{enint'}} \cdot \underset{(130\times130)}{(\mathbf{I - A})^{-1}} \cdot \underset{(130\times97)}{\mathbf{FDstruct}} \# \underset{(97\times1)'}{\mathbf{FDlevel'}} \qquad (5.4)$$

where # means element by element multiplicatiopn and ' means transpose. Also this model can be used for different types of emissions as long as **emis** and **emcoef2** are opdated accordingly. All of the variables are the same as in (5.3), except for

emcoef2 which is a 1×40 aggregated version of **emcoef**.

The result of this decomposition is a 1×97 vector. Notice in this connection, that it is necessary to transpose a number of the variables.

As a test of the consequenses of aggregating before the decomposition analysis the following model was was tried as well

$$\underset{(1\times97)}{\mathbf{emis}} = (\underset{(130\times40)}{\mathbf{emcoef}} \# \underset{(130\times40)}{\mathbf{emix}} \# \underset{(130\times1)}{\mathbf{enint}} \cdot \underset{(40\times1)}{\mathbf{summa}})' \# \underset{(130\times130)}{(\mathbf{I - A})^{-1}} \cdot \underset{(130\times97)}{\mathbf{FDstruct}} \# \underset{(97\times1)'}{\mathbf{FDlevel'}} \qquad (5.5)$$

No new variables are introduced in this model. The emission coefficient variable is here the full 130×40 matrix. A parenthesis is put around the first four variables and this result is transposed to a 1×130 vector. Results of these equations can be seen in section 6.

5.3. Direct emissions from households

The decompositions presented above are only concerned with the indirect emissions caused by the final demand by industries, households and the export markets. But actually, there are quite significant direct emissions from the households that are not covered by the models above. Total emissions from households are composed in the following way

$$\mathbf{emishh} = \mathbf{emishh}^{d} + \mathbf{emishh}^{id} \qquad (5.6)$$

emishh total emission from households (could be CO_2, SO_2 and NO_x)

emishh^{id} total indirect emissions (covered by formulas (12) and (13) above)

emishh^{d} total direct emissions.

The indirect emissions by households are the emissions by industries caused by the demand of households for produced goods and services. Because the model is based on an i/o model the pollution generated by production of input to those industries are also counted and the production of input to those who produce input – and so on - are counted as well. These emissions are registered under the industries, which generated them.

The direct emissions by households are generated by the use of electricity, gas and heating. The CO_2, SO_2 and NO_x generated by this consumption is actually emitted by powerplants and district heating facilities, but due to the use of the "gross energy method" these emissions are attributed to households because they are the underlying reason for this emission. On top of that the direct use of fuel for heating and petrol for private cars are registered on the account of households.

The decomposition equation is quite simple when it comes to the direct emissions

$$\underset{(1\times1)}{\mathbf{emishh}} = \underset{(5\times1)'}{\mathbf{enconshh'}} \cdot \underset{(5\times40)}{\mathbf{emixhh}} \cdot \underset{(40\times1)}{\mathbf{emcoefhh}} \quad (5.7)$$

Also this model can be used for different types of emissions as long as **emishh** and **emcoefhh** are updated accordingly. The elements in the equation are the following

emishh A scalar of emissions of either CO_2, SO_2 or NO_x from households.

enconshh Is a 5×1 vector of energy consumption. The 5 categories of private consumption are electricity, gas, fuel, district heating and petrol for private cars. The vector is a row sum version of the full 5×40 matrix. Thus, in the model (5.7) the changes in the emissions can be ascribed to the change in the size of the energy consumption through this variable.

emixhh Is a 5×40 vector of energymix. The full 5×40 matrix of energy consumption is divided by its row sums (which is actually the vector **enconshh**), so the row-sums of emixhh equal 1. This matrix represents the consumption of the 40 energy carriers per unit of total energy consumption for each of the 5 categories. So in the model (5.7) the change in emissions can change as a consequence of changes in the 5 energy consumption goods divided by 40 energy carriers.

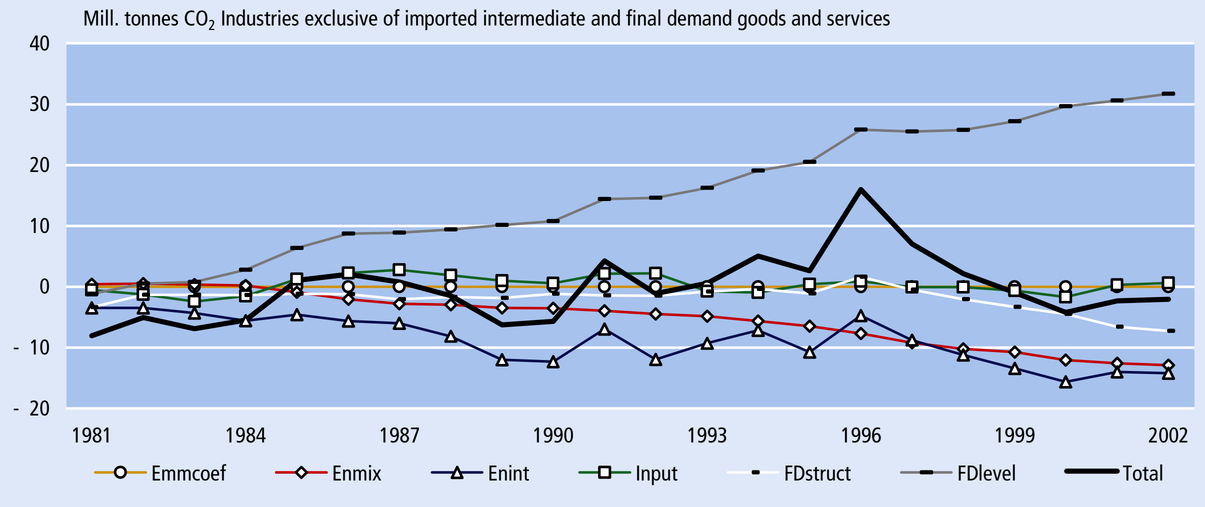
emcoefhh This is a 40×1 matrix of emission coefficients calculated as the emissions per demanded unit of energy for each of the 40 energy carriers. Through this variable the changes in emissions can be ascribed to changes in emission factors.

6. Results

As mentioned previously, total emissions in Denmark encompass the sum of emissions from all industries (including indirect emissions from households) and the direct emissions from Danish households. Firstly, we take a look at result generated with the decomposition equation (5.3) for all industries remembering that these results are the mean of the results of all *n!* decompositions .

6.1. Emissions from industries, general results

Figure 6.1 Decomposition of CO_2 emissions from all industries 1980 – 2002



The bold line in this figure indicates the total change in CO_2 emissions in Denmark from all industries as compared to the level in 1980 as it was presented in figure 1.1. It is striking that in 2002 this line is very close to zero indicating that the level of CO_2 emissions were almost exactly the same in 2002 as in 1980. This result covers a tendency to increasing emission from 1980 to 1996 and then a sharp decline from 1997 through 2000. As it is very often found in studies like this, the most significant determinant in terms of increasing the CO_2 emissions is the level of final demand. The isolated effect from final demand is an almost 30 million tonnes increase in CO_2 emissions. The peak in 1996 was due to an extraordinary large export of energy. No other determinant pulled emissions in an upward direction. Emission coefficients had absolutely no effect on this development since it remained constant through the years. Fortunately, the remaining determinants all pulled emissions down.

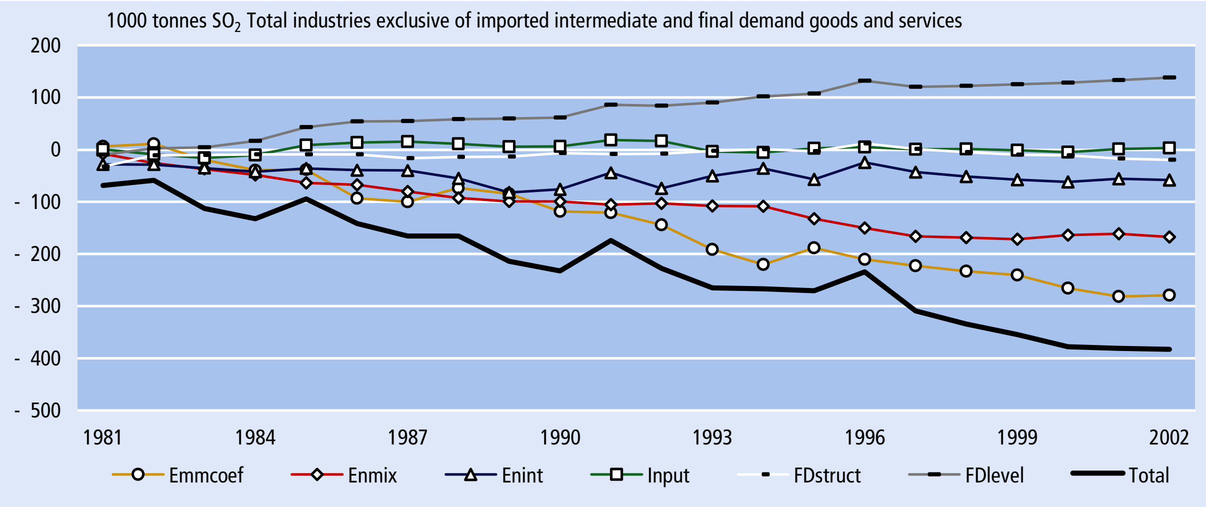
An environmentally friendly, but relatively small contribution comes from the "structure of final demand". It covers the fact that the composition of final demand has changed in the direction of goods and services that generate less emission than the previous compositions did.

The overall energy intensity in the Danish industries has decreased, meaning that the production of one unit of total output generally requires less energy input than it did the year before, because of technological development. This, off course, helps to diminish emissions. In 2002 a decrease of about 14 million tonnes can be attributed to this effect. This is also a very commonly found result in decomposition studies. It is interesting that in the years where final demand has little peaks it generates peaks in the energy intensity, probably because old, marginal and less efficient power plants come into use and play a bigger role in those years.

Also, the change in “energy mix” from 1980 to 2002 has had a favourable effect on emissions. Thus, the mix or composition of energy input has changed so that less polluting energy carriers have constituted an increasingly larger portion of the total energy input throughout the period from 1980 to 2002. This is clearly the effect of a gradually heavier reliance on natural gas and wind power as opposed to fuel oil and coal. The result of this effect is a decrease in 2002 compared to 1980 of about 13 million tonnes CO_2.

Finally, a decrease of about 7 million tonnes can be ascribed to the “structural change” in the Leontief Inverse matrix of input requirements for the industries.

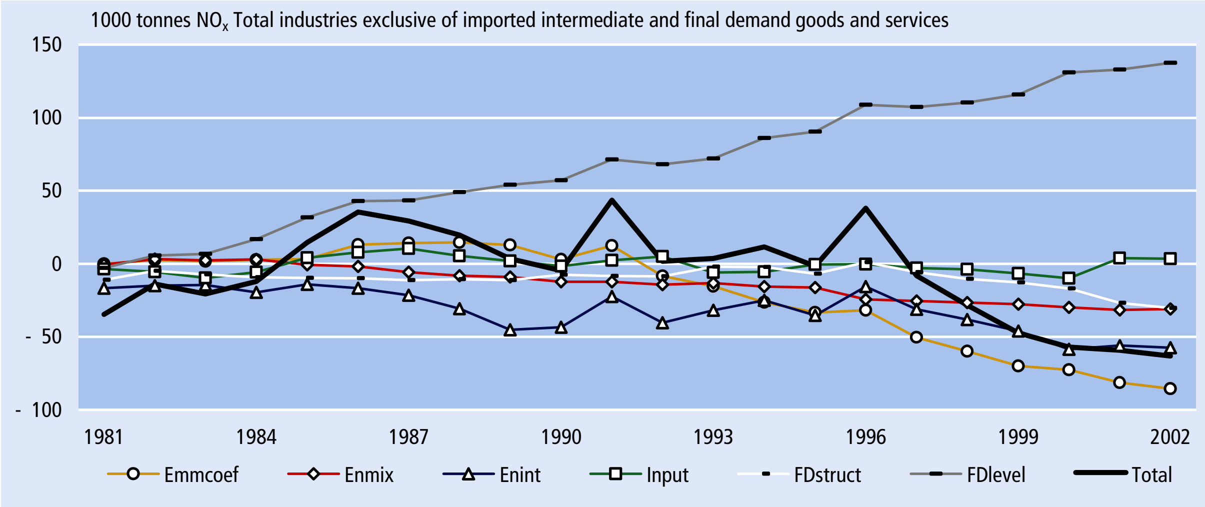
Figure 6.2 Decomposition of SO_2 emissions from all industries 1980 – 2002



When it comes to SO_2 emissions as in figure 6.2 the picture is clearer. The positive effect on SO_2 emissions from the level of final demand is not at all as dominating as in the case of CO_2 emissions. The isolated effect from final demand level is an increase of about 140,000 tonnes. However, all other determinants pull unanimously in the downward direction, thus leading to a total decrease in SO_2 emissions of almost 400,000 tonnes in 2002 compared to 1980. The best catalysts for these processes have been changes in emission coefficients and energy mix. It might seem that these two determinants are quite dependent and that they might explain part of the same effect. However, energy mix covers the extent of change from oil to coal and to some extent to natural gas and wind power, which are quite less polluting technologies. Conversely, the determinant “emission coefficient” covers the magnitude of change in improvements of e.g. sulphur content of the particular energy carriers. So these two determinants might seem closely dependent, but it is not necessarily the case. Improvements in sulphur content can happen without simultaneous changes in energy mix.

The change in energy intensity has had about the same effect as in the case of CO_2, but fortunately it is strongly dominated by the two effects mentioned above.

Figure 6.3 Decomposition of NO_x emissions from all industries 1980 – 2002



The picture for NO_x emissions is quite similar to the pictures for CO_2 and SO_2. Actually, it is much closer to the CO_2 picture than to the SO_2 picture. The total emission of NO_x from all industries has decreased by about 50,000 tonnes between 1980 and 2002. This is a much better outcome than in the CO_2 case. This is mainly due to the improvement in emission coefficients related to NO_x in the energy consumption. A main factor in this development has been the introduction of catalytic converters on motor vehicles. From approximately 1990 the level of final demand is the only determinant pulling the NO_x emissions up. From about 1994 the determinant pulling most strongly in a downward direction is the emission coefficients, but also energy intensity and energy mix helps to decrease emissions.

As one may remember from the methodological discription of the model used for decomposition, the results presented above, represent the average of 6! = 720 decomposition equations since there are 6 determinants. In order to see how well the model is doing it is obvious to try to find the minimum and the maximum values among the 720 suggestions and also to calculate the statistical standard deviations related to the means.

Here the following formula is used to calculate the standard deviation is the following

$$S_x = \sqrt{\left[\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}\right]} \qquad (6.1)$$

where *i* runs from 1 to 720, *x* is the value of the particular determinant and *x*-bar is the average of all 720 suggestions. Such statistics can be calculated for every year in the analysis. The following tables show statistics for each of the three types of pollution analysed above for the year 2002 only.

Table 6.1 Statistics on decomposition of CO_2 emissions, 2001 compared to 1980

		Emcoef	Enmix	Enint	Input	FDstruct	FDlevel
CO_2	Min	-26.2	-169.1	-125.0	-82.0	-51.5	8.3
	Max	106.9	-1.4	-1.7	-0.4	-0.4	384.6
	Mean	3.5	-13.7	-10.7	-6.5	-3.8	29.5
	Std. dev.	16.0	22.6	18.2	11.2	6.9	51.4

Table 6.2 Statistics on decomposition of SO_2 emissions, 2001 compared to 1980

		Emcoef	Enmix	Enint	Input	FDstruct	FDlevel
SO_2	Min	-3228.6	-2070.9	-783.3	-399.1	-159.6	10.6
	Max	-44.7	-6.7	-3.3	-2.2	8.0	3041.9
	Mean	-275.0	-139.4	-52.4	-25.9	-10.9	133.0
	Std. dev.	434.7	274.4	105.1	52.7	22.3	385.0

Table 6.3 Statistics on decomposition of NO_x emissions, 2001 compared to 1980

		Emcoef	Enmix	Enint	Input	FDstruct	FDlevel
NO_x	Min	-864.6	-530.3	-651.6	-243.6	-189.9	40.8
	Max	-10.9	37.9	-10.7	-5.0	-0.7	1859.5
	Mean	-71.8	-30.8	-50.8	-22.2	-11.5	131.8
	Std. dev.	116.8	73.2	90.5	35.8	24.5	240.8

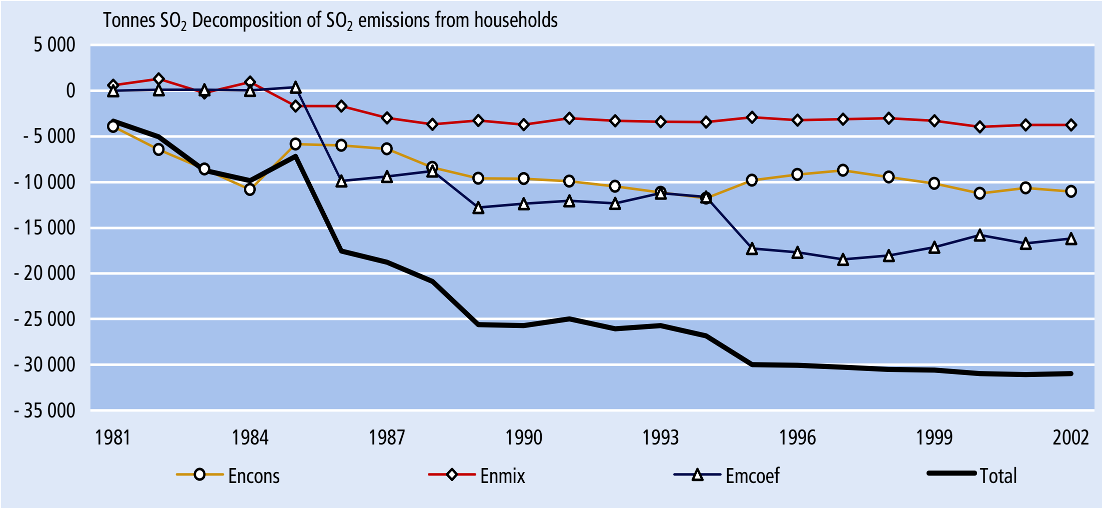
The sums of the rows of means express the total effects on the emissions in 2001. As a test on the validity of the data it is possible to refind those sums in the figures above. The statistics is a really good assurance of the danger of just picking one out of the *n!* decomposition equations as a reasonable representative of the actual values. The standard deviation numbers show that there is a huge amount of variation among the individual decompositions, and that the only way forward is to use some kind of average.

6.2. Emissions from households

Using the model (5.7) we can analyse emissions from households as the other main group of emissions. From the 40 energy carriers 5 groups of special importance for households are extracted. They cover the direct emissions by households. The decrease from 1980 to 2002 in CO_2 emissions directly from Danish households amounted to about 2.5 million tonnes. This is actually a larger decrease than from all industries in total. The two determinants "emission coefficients" and "energy mix" had close to no influence on this result. Thus, all of the decrease in emission has come from a similar decrease in energy consumption in general. Use of energy in Danish households has become more efficient since the beginning of the period under analysis.

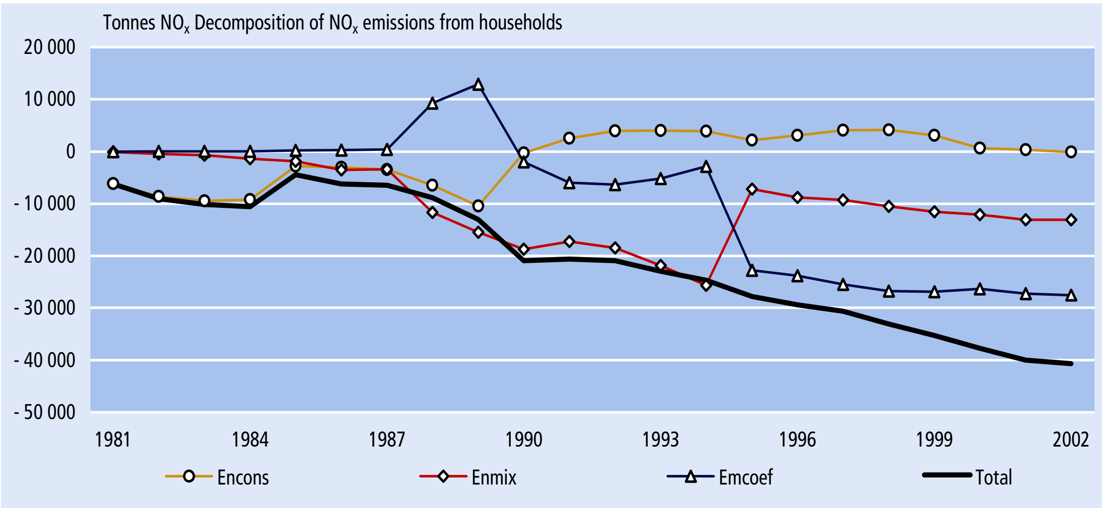
In the two figures below the results for the decomposition of SO_2 and NO_x emissions are shown. As it can be seen SO_2 emissions from households have decreased by approximately 30,000 tonnes which is fine, but not as good a result as for CO_2 when compared to the almost 400,000 tonnes decrease brought about by the industries. The energy mix does not seem to be a very effective determinant, but emission coefficients have meant more in the process, responsible for about half of the decline in emissions. Again we can refer to the lowering of the sulphur content in coal and fuel oil used to generate electricity and district heating, as the main explanation. Naturally, the decreasing size of the energy consumption has helped decrease emissions.

Figure 6.4 Decomposition of SO_2 emissions from households 1980 - 2002



In the case of NO_x we see a different story.

Figure 6.5 Decomposition of NO_x emissions from households 1980 - 2002



The 40,000 tonnes decrease in emissions from households is fully competitive with the 50,000 tonnes decrease jointly brought about by all Danish industries. The absolute main explanation is the decrease in emission coefficients. Again, one of the primary explanations is that catalytic converters on motor vehicles operated by households, have been installed since about 1990. The two determinants energy consumption and energy mix mean very little in this account.

6.3. Disaggregated results for industries

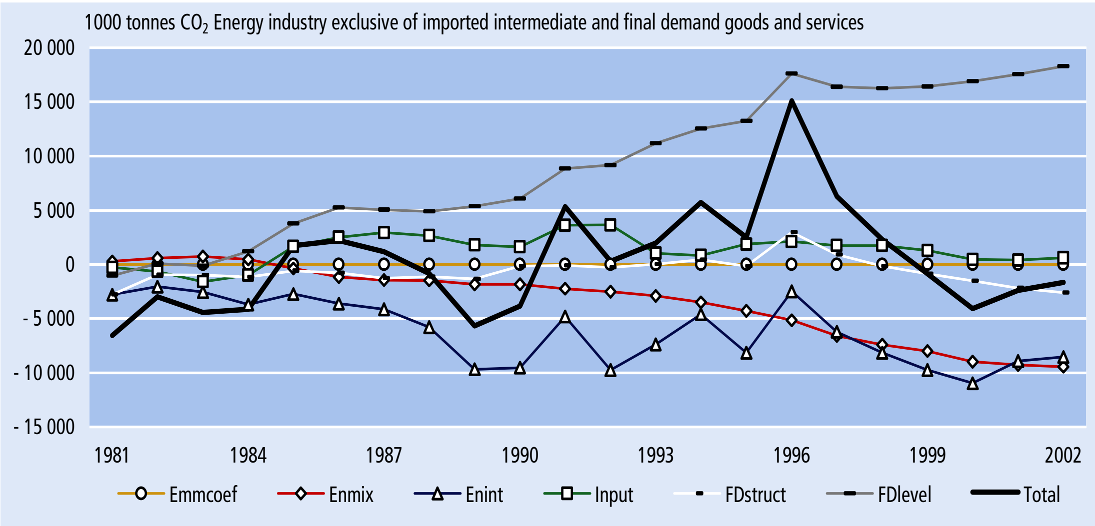
With the general tendencies from the overall economy represented by industries and households in place, we can now turn to results for more disaggregated sectors of the economy. In the process of running the decomposition analyses for this report and looking at the results, it became clear that the more disaggregated the data are, the larger effects of the decomposition analysis. Thus, the more aggregated input data for the decomposition analysis is, the more information is lost. Therefore it is preferable to do the analysis on a level as detailed as possible, and then aggregate the results.

Analyses carried out with the model (5.3) result in a (130H1) industry by emission vector, but we are not interested in the details of all 130 industries. Therefore results are aggregated to the following groups, which are actually a further aggregation of the BR9 grouping in the Danish national accounts.

1. **BR9: 1** Agriculture etc.
   (Agriculture, fishery, horticulture, mining and extraction of crude petroleum, natural gas and minerals)
2. **BR9: 2** Manufacturing industries
   (manufacturing industries other than energy supply BR9: 2)
3. **BR9: 3** Electricity, gas, district heating and water supply
4. **BR9: 6** Transport, storage and communication (ground transportation, air- and water transportation etc.)
5. **BR9: 4+5+7+8** Other industries (construction and services)

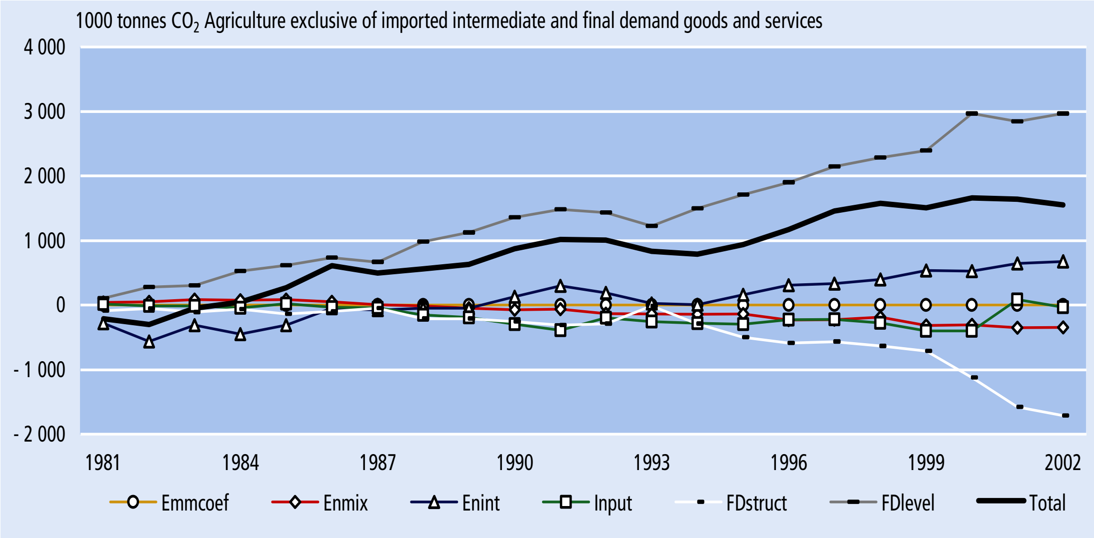
The largest emissions come from the Electricity, gas, district heating and water supply, group number 3, so it is obvious to take a closer look at this group.

Figure 6.6 Decomposition of CO_2 emissions from the Electricity, gas, district heating and water supply industry 1980 - 2002



This graph is quite close to the graph showing CO_2 emissions for all industries, because this group of industries is responsible for the major part of total emissions from industries. The level is just somewhat lower and there small differences between the determinants.

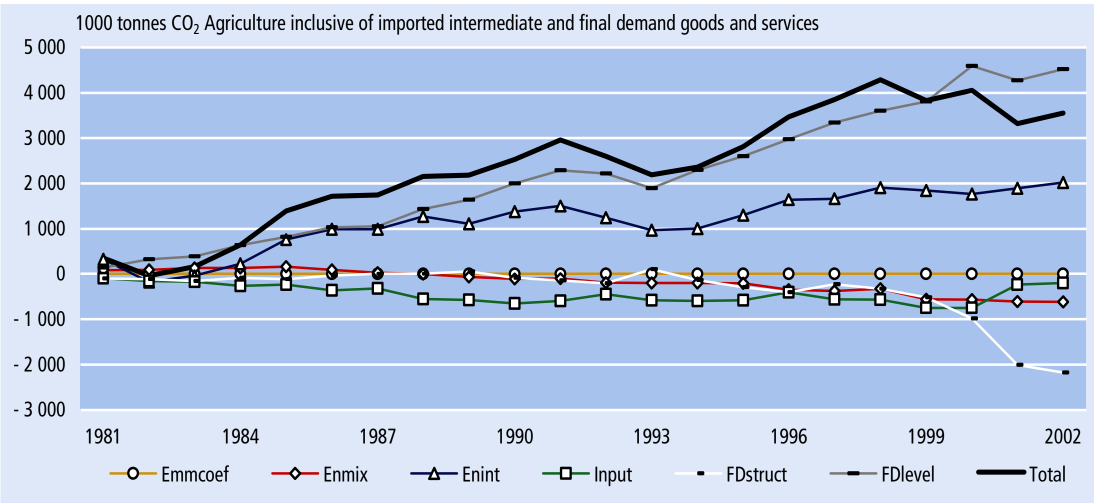
Figure 6.7 Decomposition of CO_2 emissions from agriculture 1980 - 2002



Not in every sector of the economy emissions has gone down. As shown in figure 6.7., CO_2 emissions in the sector "Agriculture etc." has gone up. In this case the level of final demand pulls a 1,5 million tonnes CO_2 from this sector in the period 1980-2002. The best counterweight is final demand structure, but it pulls far from enough in the opposite direction. If one looks behind this development it is revealed that Agriculture itself is not the big sinner here. Extraction of crude petroleum and gas in the North Sea is the generator of this development. Input of energy in the production process has increased much faster than output.

This can be seen even more clearly if we take a look at the **global** emissions generated by Danish final demand.

Figure 6.8 Decomposition of global CO_2 emissions from agriculture including imports 1980 - 2002

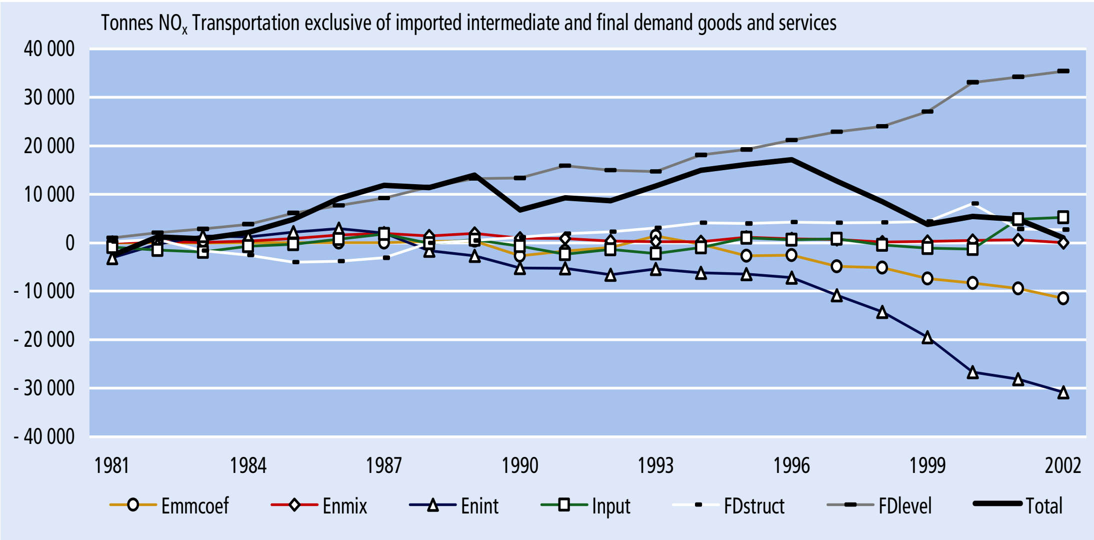


The increase in emissions in 2002 compared to 1980 is not 1,5 millions tonnes now, but almost 4 million tonnes. The level of final demand pulls a little more, but especially the contribution from the energy intensity is remarkable.

6.4. Final demand

Now let us turn to the other group of decomposition analysis based on equation (5.4). Here the matrices are transposed before they are multiplied, so the result is emission by final demand component instead of emission by industry. As an example of the results please take a look at the figure below

Figure 6.9 Decomposition of NO_x emissions from the transportation services 1980 - 2002



Note that until 1995 the development in NO_x emissions followed the development in the level of final demand very closely, because there was no other influence. However, after that date the energy intensity in transportation services has improved quite a lot. Despite the ever-increasing final demand level, this improvement has almost managed to bring NO_x emissions down to the 1980 level. In recent years, also the emission coefficients have helped to bring pollution down.

References

Ang, B. W. & Zhang, F. Q. (2000): A survey of index decomposition analysis in energy and environmental studies. *Energy 25*, pp. 1149-1176.

Bacharach, M. (1970): Biproportional matrices and input-output change. *Cambridge [Eng.], University Press*.

Betts, J. R. (1989): Two Exact, Non-Arbitrary and General Methods of Decomposing Temporal Change: *Economics Letters*, 30, pp. 151-156.

Carter, A. (1970): *Structural Change in the American Economy*. Cambridge, MA, Harvard University Press.

Casler, S. & Rose, A (1998): Carbon Dioxide Emissions in the U.S Economy. *Environmental and Ressource Economics* 11 (3-4), pp. 349-363.

Chen, C. Y. & Rose, A. (1990): A structural decomposition analysis of energy demand in Taiwan, *The Energy Journal*, 11, pp. 127-146.

De Haan, M. (2001): A Structural Decomposition Analysis of Pollution in the Netherlands. *Economic Systems research*, Vol. 13, No. 2.

Dietzenbacher, E. & Los, B. (1998): Structural Decomposition Techniques: Sense and Sensitivity. *Economic Systems Research*, 10, pp. 307-323.

Dietzenbacher, E. & Los, B. (2000): Structural decomposition analyses with dependent determinants. *Economic Systems Researc*h, 12, pp. 497-514.

Fujimagari, D. (1989): The Sources and Change in Canadian Industry Output. *Economic Systems Research*, Vol. 1, No.2.

Hoekstra, R. & Van der Berg, J. C. J. M. (2002): Structural Decomposition Analysis of Physical Flows in the Economy. *Environmental and Ressource Economics*, 23, pp. 357-378.

Hoekstra, R. & Van der Berg, J. C. J. M. (2003): Comparing structural and index decomposition analysis. *Energy economics*, 25, pp. 39-64.

Jackson R.W., & Murray, A.T. (2002): Alternate Formulations for Updating Input-Output Matrices. Research paper*, 2002-9*. West Virginia University.

Jakobsen, H. K. (2000): Energy Demand, Structural Change and Trade A Decompoosition Analysis of the Danish Manufacturing Industry. *Economic System Research. Vol. 12, No. 3.*

Jensen, H. Vadmand & Pedersen, O. Gravgård: Danish NAMEA 1980-92, Statistics Denmark.

Leontief, W. & Ford, D. (1972): Air Pollution and Economic Structure: Empirical Results of Input-Output Computations. In *A.* Brody and A. Carter, eds., *Input-Output Techniques.* North Holland, Amsterdam, The Netherlands.

Lin, X. & Polenske, K. R. (1995): Input-Output Anatomy of China's Energy Use Changes in the1980s. *Economic Systems Research*, 7, pp. 67-84.

Liu, X. Q., Ang, B. W., & Ong, H. L. (1992): The Application of the Divisia index to the Decomposition of Changes in Industrial Energy Consumption. *The Energy Journal* 13 (4), pp. 161-177.

Munksgård, J., Pedersen, K. A., & Wier, M. (1998): *Miljøeffekter af privat forbrug.* (Environmental Effects of Private Consumption). AKF Rapport. Amternes og Kommunernes Forskningsinstitut. (Report in Danish by the Institute of Local Government Studies - Denmark).

Ploeger, E. (1984): The Effects of Structural Changes on Danish Energy Consumption, 5th IIASA Task Force Meeting on input-output modeling. Laxenburg, Austria.

Rose, A. & S. Casler (1996): Input-Output Structural Decomposition Analysis: A Critical Appraisal, *Economic Systems Research*, 8, 33-62.

Rose, A., & Chen, C. Y.(1991): Sources of change in energy use in the U.S. Economy, 1972-1982. *Res. Energy 13*, 1-21.

SEEA (2003): Handbook of National Accounting – Integrated Environmental and Economic Accounting. ST/ESA/STAT/SER.F/Rev.1 (Final Draft) United Nations et. al.

Seibel, Steffen (2003): Decomposition Analysis of Carbon Dioxide Emission Changes in Germany - Conceptual Framework and Empirical Results. European Commission. Working Papers and Studies.

Skolka, J. (1989): Input-output structural decomposition analysis for Austria. *Journal of Policy Modeling*, 11, pp.. 45-66.

Wier, M. (1998): Sources of Change in Emissions from Energy: A Structural Decomposition Analysis. *Economic Systems Research*, 10, pp. 99-111.

Zheng, Y. (2000): Sources of China's Energy Use Changes in the 1990's. 13th International Conference on Input-Output techniques, Macerata, Italy.