Aggregation in Macroeconomic Models: An Empirical Input-Output Approach

Asger Olsen[1]

August, 1999

Abstract

The aggregation level of industries in the Danish macroeconomic model ADAM is examined using a new indicator of aggregation bias. The indicator is decomposed into contributions from the original industries, thereby clearly identifying the aggregation problems which caused the 6 industry groups of the older versions of ADAM to be disaggregated into the current 19 groups.

An aggregation key minimizing the new bias indicator is found: From the microlevel of 64 industries, 18 "optimal" industry groups are formed through "clustering"; these groups are very similar to the current ADAM groups.

Altogether, the conclusions based on the new indicator closely resemble those reached through years of practical experience.

JEL classification: D57.

Key words: Aggregation, input-output, macroeconomic model, clustering

---

[1]Macroeconomic Model Unit, Statistics Denmark, Sejrøgade 11, DK-2100 Copenhagen, Denmark. Phone (+45) 39 17 32 01, Fax (+45) 39 17 39 99, E-mail: jao@dst.dk.

This article summarizes research done at Unibank, Unidanmark, supported by the Danish Social Science Research Council, grant no 14-5059

Contents

1. Introduction

This article deals with the aggregation level of industries in the danish model ADAM, demonstrating the usefulness of an indicator of aggregation bias proposed by the author.[2] The analysis is carried out using Danish data, since the purpose is to show how the aggregation bias indicator easily identifies the same problems which were in fact gradually revealed through the regular use of the model over many years; the "Annual Danish Aggregate Model" has been used by the government for economic policy analysis, budgeting and forecasting purposes for more than 20 years. The model is in the econometric tradition of Tinbergen and Klein, but it contains an integrated, structural form input-output system for determination of production and prices.

The Danish system of national accounts provides a yearly series of fully specified input-output tables from 1966 onwards at a detailed level of 117 industries.[3] The first version of the ADAM model taking full account of the yearly input-output tables was the version of September 1979, having 6 main industry branches:

Table 1. The industry branches of ADAM, September 1979 version

Code	Description	Production 1980 (mill. DKK)
a	*Agriculture etc.*	39087
n	*Manufacturing*	210694
b	*Construction*	52148
h	*Dwellings*	38956
o	*Production of government services*	105241
q	*Production of services etc.*	177030

[2]The bias indicator was developed in Olsen(1993). A survey of the ADAM model is given in Dam(1986); the present model is documented in detail on our website, *www.dst.dk/adam.*

[3]These figures, as all data in this paper, are taken from the SNA68 version of the national accounts. The most recent data, on SNA93 basis, have not yet been published before 1988.

Through a static, open input-output model, these 6 main branches provided the link between 14 types of primary inputs and 27 categories of final demands (the numbers of primary inputs and final demands include 11 components of imports and 7 components of exports, respectively, with commodities broadly by 1-digit SITC).

Through the day-to-day use of the model for government purposes, it became clear that a disaggregation of the 6 branches was desirable. The main problems caused by the high aggregation level of industries were identified by the users as:[4]

- Problems getting the highly energy-intensive final demands, e.g. consumption of gasoline and fuel, to draw from the correct categories of imports. Such problems were caused by the inclusion of oil refineries and electrical power stations in the aggregated manufacturing sector

- Problems getting the large price increases on imported oil to push up the prices of the correct components of final demand, e.g. the consumption of gasoline and fuel. Clearly, these symptoms are duals of those mentioned above, having the same basic cause

- Problems in the determination of the import of services. The Danish structure in this field is very particular, due to the huge merchant marine of our country. This marine provides substantial exports of transport services, accompanied by derived service imports (mainly ship repair in foreign harbors etc). This structure was heavily distorted through the inclusion of ocean transport services in the aggregated branch of service producers, which is mainly domestic.

[4]See Danmarks Statistik(1985), chapter 9.

- Problems in the determination of margins of trade in an input-output context. The users did not consider the standard price model well suited for the short-termed determination of trade margins.

- Problems linking production and exports of agricultural products, since most agricultural exports pass through the food manufacturing industries, which contribute only little to the product price. Since the food manufacturing industries were included in the aggregated branch of manufacturing, the links from agriculture to agricultural exports were heavily distorted, concerning both prices and quantities. Furthermore, the standard input-output model was not believed to be an adequate description of the production terms of agriculture, in particular for EEC countries like Denmark.

After thorough considerations by a working group, the 6 old ADAM branches were disaggregated into 19 new ADAM branches. For the sake of the current use of the model for budgeting and other government purposes, the ADAM branches had to be simple aggregates of the 64 industries of the preliminary versions of the national accounts.

The 19 new branches chosen for the model versions of November 1984 version and later, and their definition from the 64 "detailed industries" are listed in table 4, columns 1 through 3.[5] We will re-evaluate this choice in the light of the results developed in Olsen(1993), allowing for the obvious risk of hindsight.

[5]The considerations were reported by the comittee in Danmarks Statistik(1982). A detailed summary is in Danmarks Statistik(1985), chapter 9. An english summary of the applied principles for aggregation is contained in Olsen(1985). The 64 detailed industries are listed in appendix B.

2. Evaluating the old ADAM branches

A new indicator of "aggregation bias" was proposed in Olsen(1993). In the case of aggregation of *industries* into a smaller number of *main branches* of production, the contribution of each detailed industry to the bias indicator is

$$\beta_i = (p_i - p_h^*) \cdot (g_i - w_i g_h^*) \qquad i=1,...,n \quad \textbf{(1)}$$

where

- p_i is the price index of production in industry i
- p_h^* is the price index of production in the main branch h, to which industry i belongs
- g_i is the production in industry i
- g_h^* is the production in the main branch h, to which industry i belongs
- $w_i = g_{io}/g_{ho}^*$ is industry i's share of the main branch production in the base year of the input-output table

(the n-vectors of β_i's, g_i's and p_i's are denoted β, **g** and **p**, respectively). The indicator should be interpreted as follows: Small absolute values of the price term $|p_i - p_h^*|$ for industries in group h indicate that the industries in the group have similar input structures and, therefore, similar costs and prices. Likewise, small absolute values of the quantity term $|g_i - w_i g_h^*|$ indicate that the industries in the group have similar demand structures and, therefore, roughly proportional productions. Thus, a small absolute value of the total contribution to (1) from group h requires either supply side similarity or demand side similarity of the industries in the group (but not necessarily both).

We will use the sum of absolute values, $\sum_i |\beta_i|$, as an overall indicator of the quality of an aggregation.[6]

In order to analyze the internal consistency of the aggregation, the prices in (1) should be values calculated from the detailed and aggregated input-output price models, using a variety of relevant exogenous primary input prices. Likewise, the quantities should be predicted values from the detailed and aggregated input-output quantity models, using a variety of relevant exogenous final demand quantities. In this article, the values of β_i are calculated from the 64-industry Danish input-output table of 1980, using actual values of the exogenous variables covering the period 1966-1988.[7] We will use the *mean absolute aggregation bias* of the period, defined by

$$ma(\beta) = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}|\beta_{it}| \qquad \textbf{(2)}$$

If $ma(\beta)=0$, then the aggregation is perfect for the range of values of exogenous variables covered by the period 1966-88 (unfortunately, the converse statement is not always true, due to the possibility of "mixed aggregates"), see Olsen(1993,1999). The $ma(\beta)$ indicator and some of the large individual contributions to it is shown in table 2.

[6]Olsen(1993) uses the "root mean square" $rms(\beta)$ as the overall bias indicator. While the *rms* measure is simpler for theoretical purposes, the $\sum_i |\beta_i|$ measure has a number of practical advantages, including additivity of the individual contributions.

[7]Since the official input-output tables at constant prices of Danmarks Statistik do not include a splitting of gross value added at factor prices into components of *compensation of employees* and *gross operating surplus*, a common index of *gross value added per output unit* was used in the computations of predicted prices. Thus, differences in the "wage share" of industries are not taken into account in the evaluation of aggregations; only differences in the "value added share" counts.

Table 2. Contributions to aggregation bias, old ADAM branches

no	Part of	Industry	*ma(β)*	Cause
25	*n*	*Petroleum refineries*	273.6	price
48	*q*	*Ocean transport*	58.0	price
40	*n*	*District heating*	45.5	both
38	*n*	*Supply of electricity*	36.7	both
35	*n*	*Shipyards etc*	35.6	quantity
33	*n*	*Manuf. of machinery*	24,9	quantity
9	*n*	*Slaughtering etc.*	24.4	quantity
		Other industries, average	5.2	
		Total bias	794.7	

Note: The contributions are computed by substitution of *predicted* prices and quantities for 1966-88 in (1) using the aggregation key of the 6 old ADAM branches.

In addition to the industries of table 2, some small industries could be mentioned, having a large bias expressed in percentage of production. Such industries may not contribute much to the total bias, but they still clearly do not fit in their groups. The most significant examples are numbers *39* (gasworks and distribution), *26* (manufacturing of asphalt etc.) and *7* (extraction of coal, oil and gas).

Clearly, four out of the five problems originally stressed by the model users are identified in table 2. Industries nos *25*, *35* and *48* have by now become independent ADAM branches, see table 4. The remaining industries listed in table 2 have become dominating parts of one new ADAM branch each: No *9* (slaughtering) is dominating the *nf* group (manufacturing of foods), no *33* (manufacturing of non-electrical machinery) is dominating the *nm*-group (manufacturing of machinery), while nos *38*,*39* and *40* constitute the *ne*-group (public supply of energy).

However, the fifth problem stressed by the users, concerning the determination of trade margins, is not identified in table 2. This is no defect of the aggregation bias

indicator, since the users demanded the separation of the trade margin sector from the other services because the *detailed* input-output model was supposed to be an inadequate description of this industry. Of course, an aggregation bias indicator is not useful for judging the validity of the detailed model. However, the user argument for cutting out the trade sector from the other service industries is still a relevant one indeed.

We conclude that the aggregation bias is only a part of the information relevant for the evaluation of aggregations in practice. It is equally important to identify those industries not adequately described by the detailed model, since such industries should not be put in the same group as more "well-behaved" industries.

Fortunately, the bias indicator (2) can give us hints of this identification of "misbehaving" industries too, if we form the indicator using the *actual* quantities **g** and prices **p**, rather than the predicted variables from the input-output models. Now, if the models did hold exactly, this would yield precisely the same figures as those of table 2. But the detailed models do not hold exactly, and thus some new industries enter the list of heavy contributors to the total bias. The newcomers are those industries for which the detailed model is most seriously inadequate, compared to the other industries in the same group. The results are shown in table 3.

Table 3. Contributions to aggregation bias, old ADAM branches

No	Part of	Industry	*ma(β)*	cause
25	*n*	*Petroleum refineries*	623.9	both
9	*n*	*Slaughtering*	285.5	(quantity)
52	*q*	*Communication*	153.4	price
47	*q*	*Misc. land transport*	116.7	both
62	*q*	*Domestic services*	94.5	price
44	*q*	*Retail trade*	89.4	(quantity)
7	*a*	*extraction of oil, gas and coal*	88.3	price
		Other industries, average	17.6	
		Total bias	2456.9	

Note: The contributions are computed by substitution of detailed *actual* prices and quantities for 1966-88 in (1) using the aggregation key of the 6 old ADAM branches.

First of all, we note that the contributions, in general, are about three times larger in table 3 than in table 2. Thus, unfortunately, the detailed model is not too precise in general. However, this should be expected, since the vectors of actual quantities and prices are free to vary in 64 dimensions, whereas the variation of the predicted quantities **g** and prices **p** is limited to 27 and 18 dimensions respectively (namely, the number of categories of final demands and primary inputs used for the predictions). It is hardly surprising that we suffer a substantial loss of prediction power using such a limited number of exogenous variables. In any case, the newcomers in the table are those detailed industries causing significant trouble in their groups due to a serious lack of conformity to the detailed model.

Table 3 confirms the need to cut out petroleum refineries and the slaughtering industry, which has already been identified in table 2. The newcomers in table 3 are all inadequately described by the detailed model: Industry no *52* (communications) has an atypical price formation in its group, which may be due to the predominantly public ownership. Industry no *47* (misc. land transport) is atypical with respect to

both quantities and prices. Both of these industries are included in the new *qt* branch of ADAM (misc. transport services).[8] Industry no *44* (retail trade) differs significantly from other service industries, in particular concerning quantities. The industries nos *62* (domestic services) and *7* (extraction of oil, gas and coal) are so atypical that they contribute significantly to the overall indicator, despite their very small share of total production.

The figures of table 3 conclude the evaluation of the disaggregation of ADAM branches by identifying the fifth and last of the originally stated problems with the use of the old 6-branch key. However, the table shows that a more detailed treatment of the transport and communication industries is equally necessary. The only industry of tables 2 and 3 which has *not* been given special treatment in the new ADAM key is the very small industry no *62* (domestic services).

Altogether, the aggregation bias indicator of (2) proves to be a surprisingly useful tool for easy and automatic identification of the problems otherwise revealed by the daily use of the model. However, it should be pointed out that the differences in the bias contributions are small, with a few noticeable exceptions. Only the contributions from petroleum refineries, slaughtering, the public energy supply industries and some transportation industries are so large that those industries must have special treatment.

The high volatility of energy prices seems to be a major cause of aggregation bias as measured in (2), since most of the industries displayed in tables 2 and 3 are heavy energy users (including the energy converting industries). Traditional aggregation bias measures, which focus solely on input coefficient similarity, would attach much less weight to energy; however, the user arguments previously listed complies better

[8]The committee report of Danmarks Statistik(1982) points out a division of the transport sector according to public or private ownership as the most urgent need for revision of the new grouping of industries in ADAM

with (2). Another difference is that, in addition, (2) provides information of demand side similarity leading to the conclusion that industries such as slaughtering and production of investment goods (machinery and ships) should be separated from the aggregated branch of manufacturing; though quite reasonable, such suggestions could never emerge from traditional bias measures.

3. An optimal aggregation key

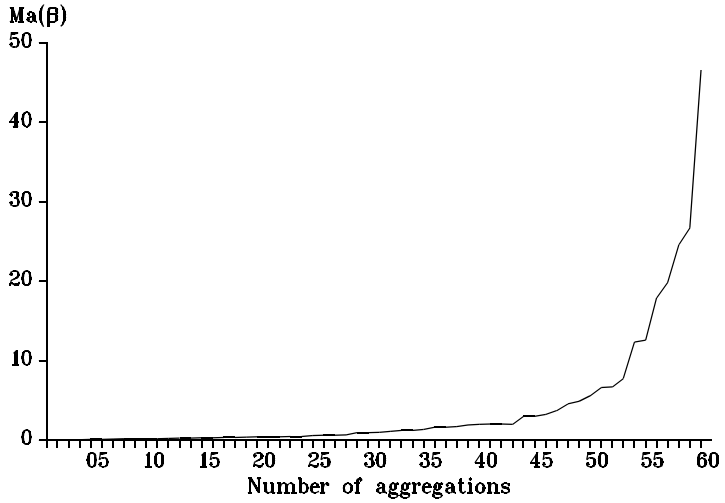
The actual disaggregation of ADAM went even further than the users suggested, since it involved the separation of the additional branches *nn* (manufacturing of beverages and tobacco), *nb* (manufacturing of construction materials), *nk* (manufacturing of chemical products) and *qf* (financial institutions). The separation of such branches is difficult to justify on the basis of the aggregation biases of tables 2 and 3. Furthermore, it is not clear from the tables whether the particular industries pointed out should be accompanied by a number of "satellite industries" in the new group.

In order to analyze these problems further, an attempt has been made to find an "optimal" aggregation key by means of *clustering*. The method chosen was "*progressive pairwise merger*", which can be outlined as follows: Starting from the most detailed level of 64 industries we simply aggregate the two industries causing the smallest aggregation bias according to (2).[9] The new group is treated in line with the remaining industries, and the procedure is repeated until a suitable number of industries is aggregated. In other words, the method forms a hierarchical structure of pairwise aggregations, each minimizing the increase in total aggregation bias.

[9]See e.g. Fisher(1969). For computational simplicity, we can use the approximations $p_h^*=\sum_{i\in h} w_i p_i$ and $g_h^*=\sum_{i\in h} g_i$ in (1), since these formulae hold in the case of perfect aggregation, see Olsen(1993,1999).

The suitable number of industry groups depends on the preferences of the model builder. Some users may want no aggregation bias at all, preferring to accept the whole burden of modelling the detailed industries. Other users may want to save modelling effort at the expense of some acceptance of aggregation bias. However, some general observations can be made: Since the variations of predicted quantities **g** and prices **p** are restricted to 27 and 18 dimensions, respectively (namely the number of final demand components and primary inputs), there is little point in having far more than 27 aggregated industries. Thus, in the iteration process, we would expect the aggregation bias to increase slowly during the first about 30 aggregations, and then to

Figure 1. Additional bias of successive "optimal" pairwise aggregations[10]



increase more rapidly for each additional aggregation. These findings are roughly supported by the actual calculations, as shown in figure 1.

[10]The figure shows the bias of the first 60 aggregations only, since the biases of aggregations 61-63 would completely dominate. The additional bias of aggregation no. 63, which is *petroleum refineries* and the rest of the economy, amounts to 314!

Table 4. Aggregation key of ADAM industries and the "optimal" key

Code	Description	ADAM, current key	"optimal" key
a	*Agriculture etc.*	*1-4,6*	*1,4, 18,20,27,29*
e	*Extraction of oil and gas*	*7*	
qs	*Ocean transport etc.*	*48*	*7,48*
ng	*Petroleum refineries*	*25*	*25*
ne	*Public energy supply*	*38-40*	*38*
			39,40 ,26
nf	*Manufacturing of food*	*9-11*	*9-11, 6,31*
nn	*M. of beverage and tobacco*	*12,13*	-
nb	*M. of construction materials*	*5,8,18,26,30*	*30*
nk	*M. of chemicals*	*23,24,27,28,37*	*23,24*
nt	*Shipyards etc.*	*35*	*35*
nm	*M. of machines*	*31-34,36*	*32-34,36*
nq	*M. of miscallaneous*	*14-17,19-22,29*	*+14-17,19*
			+13,28,37
b	*Construction*	*42*	*42*
qh	*Trade*	*43,44*	-
qf	*Financial institutions*	*53,54*	*53*
qt	*Misc. transport*	*46,47 49-52*	*46*
			47,2,8
			50
qq	*Misc. services*	*41,45,56-63*	*41,45,56-63*
			+43
			+49,51,52
			+54
			+3,5,12,21,22
h	*Dwellings*	*55*	*55*
o	*Public services*	*64*	*64, 44*

Note: The first letter in the industry code is one of the original 6 industry codes, the other letter is a suffix denoting a specific subgroup.

In table 4, the current main branches of ADAM are compared with the 18-group level of the "optimal" aggregation key. The two aggregation keys are clearly comparable, though there are, of course, differences.

First of all, we note that the ADAM branches *ng*, *nt*, *b* and *h* are immediately pointed out by the "optimal" key. Likewise, the ADAM branches *e*, *qs*, *ne*, *nf*, *qf*, *o* and partly

qh emerge from the optimal key after a rough correction for special conditions.[11] However, the "optimal" key suggests a further separation of electric power stations (no *38*) from the *ne* branch of public energy supplies. In any case, the real discrepancies are found in the remaining groups *a*, *nn*, *nb*, *nk*, *nm*, *nq*, *qt*, *qq* and partly *qh*.

Within the manufacturing industries the "optimal" key suggests an aggregation of manufacturing of machines (*nm*) and misc. manufacturing (*nq*), while it suggests a narrower definition of the branches of chemical manufacturing and manufacturing of construction materials. The "optimal" key lacks a counterpart to the branch of beverage and tobacco manufacturing which, however, is difficult to put in any other group.

Within the service industries the "optimal" key suggests a very broad group of misc. service producing industries balanced by a much more detailed description of the transport industries: Railway and bus transport (no *46*), misc. land transport (no *47*) and air transport (no *50*) are suggested three independent branches (we recall that, in addition, industry no *47* is not adequately described by the detailed model, see table 3). The suggestion is due to the higher energy cost share and export orientation in the transportation service industries.

---

[11]Extraction of energy (*7*) and ocean transport (*48*) are separated, since they are not described adequately by the detailed model, see table 3. The same argument applies to production of public services (*64*) and retail trade (*44*) which, in addition, must be separated on the grounds of different ownership alone (the proposed aggregation is due to the similar input structure of the two industries, both receiving almost exclusively gross value added). Fishing (*6*) and basic metal industries (*31*) are separated from the food industry (they are not adequately described by the detailed model). The asphalt industry (*26*) is removed from public energy supply due to the different ownerships (the proposed aggregation of these sectors is obviously due to the common dependency on oil). Insurance institutions (*54*) are moved to the *qf* branch in accordance with international standards (however, in the December 1982 version of ADAM the industries *53* and *54* were in fact placed in the same groups as in the "optimal" key).

Some proposed aggregations are crossing the traditional borderline between the three main groups of extractive, manufacturing and service industries. The "optimal" key suggests an aggregation of agriculture with a number of basic "light" industries, namely manufacturing of wood products (no *18*), manufacturing of paper products (no *20*), manufacturing of rubber products (no *27*) and manufacturing of glass products (no *29*). The reason for this somewhat odd proposal is found on the supply side, since all these industries are heavily dependent on the imports of raw materials (SITC 2), which are characterized by price increases well below the average. Though we would never put these industries in the same group as agriculture, the suggestion is quite reasonable and points strongly towards an independent branch of "basic light industries".

Other manufacturing industries are suggested to be aggregated into the branch of misc. service producers, namely breweries (no *12*), printing (no *21*) and publishing (no *22*). The reason for this proposed aggregation is double sided. On the demand side, the suggestion is due to the links to restaurants/hotels, advertising and education, and to a low export orientation. On the supply side, the industries in question require less imports than typical manufacturing industries. Once again, though we would never aggregate service industries and manufacturing industries, the suggestion is perfectly reasonable and points towards an independent "home market manufacturing" branch.

Finally, we must readily admit that a few aggregations of the "optimal" key have no sensible interpretation. Such aggregations are partly due to imperfections of the algorithm, and partly due to inadequacies of the detailed model, and they affect almost exclusively very small industries in the extractive (primary) sector; they are discussed in more detail in appendix A. We can easily move these small industries from their meaningless groups into more sensible ones virtually without affecting the total aggregation bias. In any case, the few odd aggregations should not prevent us

from taking advantage of the substantial and intuitively plausible advice given by the "optimal" aggregation key on the whole.

Altogether, the differences between the current ADAM aggregation key and the "optimal" key are generally small and have a straightforward interpretation. In fact, the similarity of the two keys is quite amazing, considering their very different origins.

4. Conclusions

The new indicator of aggregation bias is well suitable for the identification of the serious aggregation problems of older versions of the ADAM model, problems which were in fact gradually revealed through the daily use of the model. The method of progressive pairwise merger succeeds in establishing an "optimal" aggregation key, which is very close to the new aggregation key actually chosen for ADAM, despite the quite different methods involved; the differences of the two aggregation keys are generally easy to interpret and provides food for thought. The high volatility of energy prices is an important driving force in both keys.

However, some aggregation suggestions of the "optimal" key are clearly undesirable. These problems stress the need for an additional analysis of the adequacy of the detailed model, which is not usually part of the aggregation considerations. However, it should be a trivial finding that some industries do not conform at all with the detailed model, and that such industries should not be aggregated with more well-behaved industries.

REFERENCES

Dam, P. U.(1986) The Danish macroeconomic model ADAM: A survey, *Economic Modelling*, 3, january 1986, pp 31-52.

Danmarks Statistik(1982) Det fremtidige arbejde med ADAM (The future work with ADAM), *Rapport fra Modelgruppen* nr. 5 (*Report from the Macroeconomic Model Unit* no 5). Copenhagen.

Danmarks Statistik(1985) Input-Output systemet i ADAM (The input-output system of ADAM), *Arbejdsnotat (Working papers)*, 19. Copenhagen.

Fisher, Walther D.(1969) *Clustering and Aggregation in Economics* (Baltimore, The Johns Hopkins Press).

Olsen, Jes Asger(1985) Adaptation of Detailed Input-Output Information: Restructuring and Aggregation, *Review of Income and Wealth*, 1985, No. 4, pp 397-411.

Olsen, Jes Asger (1993) Aggregation in Input-Output Models: Prices and Quantities, *Economic Systems Research*, vol. 5, No. 3, October 1993, pp 253-275

Olsen, Jes Asger (1999) Perfect aggregation of industries in Input-Output Models, in: Erik Dietzenbacher (ed) *Input-Output models: Frontiers and extensions* (in print).

Appendix A. Undesirable aggregations

Though, in general, the method of progressive pairwise merger yields very reasonable results, it does produce some undesirable aggregations. Such aggregations may flow from at least three different sources.

In the first place, undesirable aggregations may emerge from industries that are not adequately described by the detailed model. A good example from table 4 is the proposed aggregation of *energy extraction* (no *7*) and *ocean transport* (no *48*). This suggestion is only possible because the predicted price on crude oil from the detailed model, which was used in the computations, has not much in common with the actual price. This type of undesirable aggregations cannot possibly be accounted for in an analysis confined to aggregation problems only.

Second, the undesirable aggregations may be due to a drawback of the algorithm showing up when the detailed industries are of very different magnitudes. Since very small detailed industries can never cause large contributions to the total aggregation bias, they can be put in virtually any group without really affecting the minimization criterion. This is not a drawback, but a desirable property of the bias criterion: Other things being equal, we would prefer to aggregate two small industries rather than two large industries. But the algorithm of progressive pairwise merger will tend to place the small industries in nonoptimal groups, since the small industries can always be aggregated in the early iterations due to their insignificant size alone.[12] Undesirable aggregations of this type can be avoided when better algorithms to find the true optimum are developed.

---

[12]A first attempt to correct this drawback could be to use the aggregation biases expressed in percentages of the base year production as the criterion, rather than the absolute biases. However, this criterion yields completely misleading results, since the relative aggregation bias on a small industry will always be huge if it is aggregated with a large industry. Thus, such a method tends to end up with one huge group of industries and many very small, but extremely atypical industries.

The third main reason for undesirable aggregations is possible *nonsense correlations* between the exogenous variables. Our calculations were based on the actual values of exogenous primary input prices and final demand quantitites for the period 1966-1988. Now, we cannot claim that this period spans all potentially relevant combinations of exogenous variables, and thus we cannot claim that an optimal aggregation key found on this basis is truly optimal in all relevant situations. The argument is particularly important if only a small number of observations is used for the predictions. This type of undesirable aggregations is corrected by increasing the number of combinations of exogenous variables used in the computations. One possibility is to collect the actual values for a longer period. [13]

There is little doubt that all of the three causes lead to undesirable aggregations in practice. However, the first two reasons clearly seem to dominate, since the problems tend to concentrate on quite small industries in the extractive (primary) class of industries, namely *horticulture* (no *2*), *fur farming* (no *3*), *forestry* (no *5*), *fishing* (no *6*), *energy extraction* (no *7*), *other mining* (no *8*) and *basic metal industries* (no *31*). In Denmark, the productions of these industries are commonly presumed supply constrained rather than demand constrained, due to limited natural resources.

The undesirable aggregations are annoying, but they do not invalidate the many reasonable and thought-provoking properties of the "optimal" aggregation key. We can freely rearrange the troublesome industries and recalculate the total bias using (2). Such rearrangements will hardly add much to the total bias while they are likely to leave us more satisfied.

---

[13]If the variation of the exogenous variables is specified in terms of moment matrices, see Olsen (1993), an easier solution is to reduce the off-diagonal elements of the covariance matrix. However, we should not go to the other extreme and ignore all covariations, since this will probably be more misleading than the plain use of historical data. A glance at official aggregation keys of statistical bureaux is sufficient to show that the criterion of correlated industry outputs is very important in practical aggregations.

Appendix B. The 64 industry groups of the Danish national accounts

No	Industry	Production 1980 mill. DKK
1	Agriculture	31500
2	Horticulture	2763
3	Fur farming, etc	986
4	Agricultural services	1006
5	Forestry and logging	699
6	Fishing	2835
7	Extraction of coal, oil and gas	419
8	Other mining	700
9	Slaughtering and meat processing	29879
10	Manufacture of dairy products	15148
11	Food manufacture, excl. meat and milk	20258
12	Beverage industries	4201
13	Tobacco manufacture	1118
14	Manufacture of textiles	5719
15	Manufacture of wearing apparel	3349
16	Manufacture of leather products	384
17	Manufacture of footwear	827
18	Manufacture of wood products, excl. furniture	4144
19	Manufacture of wooden furniture, etc.	4441
20	Manufacture of paper and paper products	3826
21	Printing and bookbinding	6163
22	Publishing	6394
23	Manufucture of industrial chemicals	6310
24	Manufacture of other chemical products	5711
25	Petroleum refineries	11540
26	Manufacture of asphalt and roofing materials	1207
27	Manufacture of rubber products	900
28	Manufacture of plastic products	3237
29	Manufacture of pottery, china, glass, etc.	1580
30	Other non-metallic mineral products	6099
31	Basic metal industries	3223

32	Manufacture of fabricated metal products	11164
33	Manufacture of machinery, except electrcal	20376
34	Manufacture of electrical machinery, etc.	8090
35	Manufacture of transport equipment	7742
36	Professional and measuring equipment	2633
37	Other manufacturing industries	2373
38	Electric light and power	7425
39	Steam and hot water supply	447
40	Gasworks and distribution	3386
41	Water works and supply	674
42	Construction	52150
43	Wholesale trade	35780
44	Retail trade	22220
45	Restaurants and hotels	9556
46	Railway and bus transport, etc.	4091
47	Other land transport	11680
48	Ocean and coastal water transport	13190
49	Supporting services to water transport	760
50	Air transport	4894
51	Services allied to transport, etc.	7077
52	Communication	6470
53	Financial institutions	12260
54	Insurance	2190
55	Dwellings	38960
56	Business services	19920
57	Education, market services	434
58	Health, market services	4922
59	Recreational and cultural services	3551
60	Repair of motor vehicles	8050
61	Household services	6727
62	Domestic services	810
63	Private non-profit institutions	1766
64	Producers of government services	105200