Arbeidsnotater

STATISTISK SENTRALBYRÅ

WORKING PAPERS FROM THE CENTRAL BUREAU OF STATISTICS OF NORWAY

IO 69/14 Oslo, 17. september 1969

SUBSTITUTION AND COMPLEMENTARITY EFFECTS ON INPUT-OUTPUT RATIOS

By Per Sevaldson

CONTENTS

Not for further publication. This is a working paper and its contents must not be quoted without specific permission in each case. The views expressed in this paper are not necessarily those of the Central Bureau of Statistics.

Ikke for offentliggjøring. Dette notat er et arbeidsdokument og kan siteres eller refereres bare etter spesiell tillatelse i hvert enkelt tilfelle. Synspunkter og konklusjoner kan ikke uten videre tas som uttrykk for Statistisk Sentralbyrås oppfatning.

SUBSTITUTION AND COMPLEMENTARITY EFFECTS ON INPUT-OUTPUT RATIOS

I. Causes of substitution and complementarity effects

The basic hypothesis in traditional input-output analysis is that input-output coefficients, i.e. the ratios of inputs from other sectors to total output in a given sector of production are constant and independent of the given sector's level of output. This is supposed to be the case when both inputs and outputs are measured in quantity units, usually in values at constant prices. This hypothesis also implies that there is no substitution between inputs originating from different producing sectors or from different groups of producing sectors.

In a study of the dispersion and the possible existence of trends in time series of input-output ratios for Norwegian sectors of production for the years 1949-60 in a 89 industry specification [1)] the standard computer program which was used also gave, as a by-product, the correlation coefficients between sets of input-output ratios for individual production sectors. The program could give correlation matrices for sets of up to twelve variables. Three of these variable positions were required for other variables in the study, and it was thus possible to include up to nine input-output ratios in each set. Since there were generally more than 9 input-output coefficients for each sector (Some of these were aggregates of others), the total set of input-output ratios for a sector had to be broken down into sets of no more than 9 ratios in each, and a complete correlation matrix for all input-output ratios of a sector was not obtained. Only certain blocks of such matrices were available, but these were then used to study some aspects of substitution and complementarity between inputs.

There are several possible causes for substitution effects - characterised by numerically high negative coefficients of correlation between input ratios - and complementarity - characterised by high positive coefficients of correlation between input ratios. The same mechanisms can give rise to both substitution and complementarity effects:

1) See Per Sevaldson: "The Stability of Input-Output Coefficients". Working papers from the Central Bureau of Statistics of Norway. IO 67/9, Oslo 1967. Mimeographed, also to appear in: "Proceedings of the Fourth International Conference on Input-Output Techniques" Ed.s. Carter and Brody. Forthcoming.

a) Substitution proper: the use of one or more inputs is reduced in relation to the volume of production and this is compensated for by a relative expansion in the use of one or more other inputs, with no changes in product mix. The direction of change is reversible, and is determined by price fluctuations. One could distinguish between

i) "Input substitution", where only 2 (possibly three) inputs are involved and

ii) "Process substitution", where a greater number of inputs are involved, indicating more fundamental changes in the production process.

b) Technological change: the use of one or more inputs is gradually reduced in relation to the volume of production and this is compensated for by a gradual relative expansion in the use of one or more other inputs, with no change in product mix. The direction of change is given. Technological change will not necessarily have to be gradual, but can also take effect as sudden, ineversible changes in input ratios.

c) Changes in production volume: For inputs which are not fixed proportions of output, the input ratois will change with the volume of production. If the inputs are (completely) determined by the volume of production, and if the elasticity of some inputs with respect to production are less than one and of others greater than one, there will be negative correlation between input ratios from these two groups, and there will be positive correlation between input ratios which all have elasticity greater than one or all smaller than one.

d) Changes in technology mix: Different establishments may use different technologies for production of the same output. Changes in the relative importance of the different technologies in total production, e.g. with the total volume of production may imply correlated expansions in the use of some inputs and reductions in the use of others in relation to the total volume of production.

e) Changes in product mix: When we have complex product mixes - as we must have in most sectors in an industry specification of only 89 sectors - the technologies for different products may be quite different, and changes in composition of the combined product may imply relative expansion in the use of some and contraction in the use of other inputs.

f) Specification changes: With gradual changes, e.g. improvements, in the statistical registration process, inputs, which were formerly classified under one delivering sector (e.g. Unspecified) may be more and more extensively classified under other deleviring sectors.

From the point of view of input-output analysis we are particularly interested in the stability of the input-output structure in relation to price fluctuations, that is in the extent of "substitution proper".

We could formalize the preceding arguments in the following way, which may clarify some of the points.

Let

$x_{ij}(t)$ be the quantity (measured in physical units or in value at constant prices) of input of type i (e.g. products from sector i) used by sector j in period t, and

$x_j(t)$ the production in sector j in period t measured in value at constant prices.

We have observations for $t = 1, 2, \ldots, T$.

$a_{ij}(t) = \dfrac{x_{ij}(t)}{x_j(t)}$ is then the input-output ratio or coefficient.

The basic, simple Leontief hypothesis can then be formulated as

(1) $$a_{ij}(t) = \alpha_{ij} + u_{ij}(t)$$

where

α_{ij} is a constant and

$u_{ij}(t)$ is a random disturbance which we assume to have expected value 0, to have limited variability and to be stochastically independent of other disturbances, i.e. for variations in i, j and t.

This hypothesis may then be confronted with alternative hypotheses about systematic changes in the $a_{ij}(t)$-coefficients.

All such alternative hypotheses, if we have to accept one of them, will imply the rejection of the basic hypothesis, but they will not all give us the same problems in trying to amend the basic model.

One alternative hypothesis is that of a production function of the classical type, where input proportions can be adjusted continuously in response to price fluctuations. If only two inputs were involved, we would have what we previously termed "input substitution" and we would expect to observe that the input ratio of one input would increase when the other decreased and vice versa, giving a negative correlation between the two, (at least if the production function is supposed to be homogeneous or nearly homogeneous of degree one).

If more than two inputs were involved, the relative movements of pairs of input ratios in respons to general price changes would not necessarily be closely correlated when we consider this general model. But we may consider a model where the concept of substitutability is given a more narrow definition, namely as the possibility of expanding the relative proportions of one group of inputs as a compensation for reductions in another group and vice versa. There may, (but need not), be more than one input in each group. If this model represented the structure in our production sectors, we would observe positive correlations between input ratios within the same group and negative correlations between input ratios belonging to different groups. A model like this is in a way a generalisation of the two input-substitution model, and is probably what many people have in mind when they talk of substitutability in relation to Leontief models. This model would give numerically high correlations between pairs of input-output ratios, whereas the general, continuous substitution model need not do so.

However, one must be quite careful here: If we think of a group of more than two inputs as mutually substitutable (e.g. electricity, fuel oil and coal as sources of energy), we are not assured of correlations between pairs of input-output ratios within the group: any two ratios may move in the same or in opposite directions, depending on the movements in the remaining ratio(s) in the group.

Apart from the cases where one group of inputs can simply serve as substitutes for another group, without causing substantial changes in other characteristics of a given production process, the "substitutability-model" will be appropriate in the following situation: Suppose there are two available production processes for the output of a sector, each characterized by a set of fixed input-output ratios. Suppose also that conditions in the industry, for instance regarding fixed capital structure, are such that both processes are used simultaneously, but that their proportions depend on prices. In this case the average input-output ratios will be veighted averages of the ratios in each process, and the weight in each period will be the same for all input-output ratios in the same process. Consequently the ratios will change "in step":

Let α_{ij}, α_{kj} be constant terms in the input ratios for process 1 and β_{ij}, β_{kj} be the corresponding terms for process 2, and let w_t be the proportion of total production produced by process 1 in period t. Let $v_{ij}(t)$ and $v_{kj}(t)$ be random disturbances terms, which are serially and mutually independent, as well as independent of the weights, w_t, so that we have

$$(1') \qquad a_{ij}(t) = w_t\alpha_{ij} + (1-w_t)\ \beta_{ij} + v_{ij}(t)$$

$$(2') \qquad a_{kj}(t) = w_t\alpha_{kj} + (1-w_t)\ \beta_{kj} + v_{kj}(t)$$

We will then have

$$\begin{aligned}\text{covar. } a_{ij}(t)a_{kj}(t) = {} & (\alpha_{ij}-\beta_{ij})(\alpha_{kj}-\beta_{kj})\ \frac{1}{T}\ \sum_t(w_t-\frac{1}{T}\ \sum_t w_t)^2 \\ & + (\alpha_{ij}-\beta_{ij})\frac{1}{T}\sum_t(w_t-\frac{1}{T}\sum_t w_t)(v_{kj}(t)-\frac{1}{T}\sum_t v_{kj}(t)) \\ & + (\alpha_{kj}-\beta_{kj})\frac{1}{T}\sum_t(w_t-\frac{1}{T}\sum_t w_t)(v_{ij}(t)\ \frac{1}{T}\sum_t v_{ij}(t)) \\ & + \frac{1}{T}\sum_t(v_{ij}(t) - \frac{1}{T}\sum_t v_{ij}(t))(v_{kj}(t) - \frac{1}{T}\sum_t v_{kj}(t))\end{aligned}$$

Under our assumptions the three last addends will tend to disappear, and the covariance will be dominated by the first term on the right of the equality sign. If the first two differences in this term have the same sign, the conveniance may be expected to be positive and it may be expected to be negative if these differences have opposite signs.

The situation with alternative production processes, each with fixed input-output ratios, is not so alien to input-output analysis as one might expect: If the outputs from the processes are distinguishable, there is a case for breaking up the industry into sub-industries, containing one process each, and thus again achieving a situation with stable input-output ratios. If the products are not distinguishable, we have a case for extending the simple input-output model to a model which allows the choice between alternatives processes in one or more industries. The case can be analytically handled by linear programming techniques, but the data problems will be more complicated than in ordinary input-output-analysis.

If we have to conclude that input-output ratios fluctuate in response to changes in prices for other reasons than the coexistence of alternative processes, much of the advantages of the simple input-output-theory will be lost. It is therefore important to try to ascertain the extent and the nature of the various forms of substitution.

Unfortunately, the type of substitutability which causes high correlations between pairs of input-output ratios is not the only cause of such correlations:

We have indicated several causes under the points a) to e) above: Points b) and f), gradual changes in technology or specifications will give ratios of the form

$$(2) \qquad a_{ij}(t) = \alpha_{ij} + \gamma_{ij}t + u'_{ij}(t)$$

and empirical covariances between such ratios will, with the usual assumptions about the disturbances be

$$\frac{1}{T}\sum_{t=1}^{T}(a_{ij}(t) - \frac{1}{T}\sum_{t=1}^{T} a_{ij}(t))\,(a_{kj}(t) - \frac{1}{T}\sum_{t=1}^{T} a_{kj}(t)) =$$

$$\gamma_{ij}\gamma_{kj}\frac{1}{T}\sum_{t=1}^{T}(t-\bar{t})^2 + \gamma_{ij}\frac{1}{T}\sum_{t=1}^{T}(t-\bar{t})(u_{kj}'(t) - \bar{u}_{kj}')$$

$$+ \gamma_{kj}\frac{1}{T}\sum_{t=1}^{T}(t-\bar{t})(u_{ij}'(t) - \bar{u}_{ij}') + \frac{1}{T}\sum_{t=1}^{T}(u_{ij}'(t) - \bar{u}_{ij}')(u_{kj}'(t)-\bar{u}_{kj}')$$

Here the first term must be expected to dominate under our assumptions and the correlation will be positive or negative, depending on whether γ_{ij} and γ_{kj} have equal or opposite signs.

Here we can still use traditional Leontief input-output analysis, if we can estimate the $a_{ij}(t)$ for instance on the basis of estimates of the constants α_{ij} and γ_{ij}, or even, - if the γ_{ij} are not too big - on the basis of observations of $a_{ij}(t-\theta)$, when θ can be made small enough.

Points d) and e), changes in technology mix and in product mix, will have effects similar to the effects of two coexisting and different processes for the same output, when there are only two technologies or two products. The only difference is that the changes are now not brought about as direct responses to changes in prices. If the changes are monotonic and gradual over time, these cases will be no different from points b) and f) and can be handled in the same way. If the changes are reversible and depending on the level of total production, these cases are similar to point c) which will be dealt with subsequently.

With more than two alternative processes, and when none of the above explanations of shifting proportions between the processes are relevant, the situation becomes more complex. Obviously, changes in product mix can be handled by breadking up the industry into separate sub-industries if we can obtain a sufficiently detailed breakdown of demand.

Point c), changes in production volume when there are non-proportional inputs, will imply relationships of the form

$$(3') \qquad x_{ij}(t) = \delta_{ij} + \alpha_{ij} x_j(t) + u_{ij}'(t) x_j(t)$$

where we assume that the disturbance term $w_{ij}'(t)\ x_j(t)$ is proportional to output, with the usual assumptions made for $u_{ij}'(t)$.

We will then have

$$(3) \qquad a_{ij}(t) = \alpha_{ij} + \delta_{ij} \frac{1}{x_j(t)} + u_{ij}'(t)$$

and

$$\text{covar. } a_{ij}(t) a_{kj}(t) = \delta_{ij} \delta_{kj} \frac{1}{T} \sum_t \left(\frac{1}{x_j(t)} - \frac{1}{T} \sum_t \frac{1}{x_j(t)} \right)^2 +$$

terms with either averages or covariances of disturbances, which under our assumptions must be expected to tend to vanish.

Here again, ordinary input-output analysis is applicable if we can estimate the coefficients α_{ij} and δ_{ij}, and if the variance of $u_{ij}'(t)$ is not too great.

We have seen that substitution in response to price fluctuations can take varying forms. Some of these - we have termed them "input substitution" - have the effect of causing negative or positive correlations between pairs of input-output ratios. Other forms need not have this effect. There are also other possible causes of systematic changes in input-output ratios. Also some of these causes will lead to correlations between pairs of input-output ratios. By studying the changes over time in input-output ratios, or their changes with changing levels of production we may be able to identify some of these causes.

Correlations due to technological change or spesification change is indicated when the correlated coefficients are also strongly correlated to time. Changes in production volume as the cause of correlation is indicated when the correlated coefficients are also strongly correlated to the volume of production. Changes in production volume as the cause of correlation is also indicated when there is a linear, non-proportional regression of the volume of input on the volume of output.

It is difficult to form apriori opinions about the relative importance of the various causes of substitution and complementarity effects.

II. Substitution and complementarity between inputs in general

In this part of the study we are interested in substitution and complementarity between inputs, characterised by their sectors of origion, but we are not concerned with whether the input is domestically or foreign produced. Consequently we have preferred to work with the following categories of input ratios [2].

(a) Competitive inputs combined (Sums of inputs originating from a given domestic production sector and the corresponding imported products)

(b) Norwegian, non-competitive products (i.e. products of which there is no corresponding import)

(c) Imports, non-competitive.

However, because of the grouping of the ratios into sets of 9, there were many sets which lacked some of the items of the type (a) Competitive inputs combined, but where the corresponding ratios belonging to one or both of the categories

(d) Norwegian, competitive or

(e) Imports, competitive

were included, i.e. one or both of the items which were summed in order to obtain "competitive inputs combined". In such cases the largest of these two items were used as a substitute for the corresponding competitive inputs combined-item. The frequencies of the various combinations are given in table 1.

2) For a full description of the complete set of types of input-output ratios (coefficients) computed see Op. cit. p. 6 f.

Table 1. Frequencies of combinations of input-types for which correlations between input-output ratios have been computed

Combination	Frequencies Absolute	Per cent
Competitive inputs combined/competitive inputs combined	433	49,1
Competitive inputs combined/Norwegian, non-competitive	23	2,6
Competitive inputs combined/Imports, non-competitive ..	5	0,6
Norwegian, non-competitive/Norwegian, non-competitive .	135	15,3
Norwegian, non-competitive/Imports, non-competitive ...	36	4,1
Imports, non-competitive/Imports, non-competitive	4	0,4
Sum combinations where both items, belong to the preferred types (a), b) and c))	636	72,1
Competitive inputs combined/Norwegian, competitive	4	0,4
Norwegian, non-competitive/Norwegian, competitive	100	11,3
Norwegian, non-competitive/Imports, competitive	113	12,8
Imports, non-competitive/Norwegian, competitive	9	1,0
Imports, non-competitive/Imports, competitive	21	2,4
Sum combinations where Norwegian, competitive or imports, competitive have been used instead of competitive inputs combined	247	27,9
Total	883	100,0

If we compare the number of correlations, computed in this way, with the possible number that might have been computed, if we had not been restricted by our computation program, we get the following picture:

Tabel 2. Possible and actually computed correlations

Sectors with possible number of correlations	Number of sectors	Possible number of correlations	Actually computed correlations in groups of 1-5	6-10	11-28	Total	Total in per cent of possible
1 - 3	22	50	49	.	.	49	98.0
6 - 10	27	214	55	137	.	192	89,6
15 - 28	19	386	15	135	187	337	87,4
36 and over	9	458	9	37	259	305	66,6
Total	77	1 108	128	309	446	883	79,7

We see from this that the coverage is quite good, if we disregard the fact that we to some extent have used substitutes for the coefficients of type a) (competitive inputs combined).

The distribution of the observed correlation coefficients is given in table 3.

The table also gives the hypothetical frequencies which we should expect to observe if the input-output ratios were independent and normally distributed about their expected values, i.e. if the equation

$$(1) \qquad a_{ij}(t) = \alpha_{ij} + u_{ij}(t)$$

was correct for all i, j and t, and with all the $u_{ij}(t)$ independent and normally distributed about zero. (This distribution may be computed from the t-distribution, since under the given assumptions the statistic

$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ follows the t-distribution with n-2 degrees of freedom.

r is the estimate of a correlation coefficient, the true value of which is zero. n is the number of observations on the basis of which the correlation coefficients have been computed, i.e. n = 12 in our computations).

The frequencies have been plotted in figure 1.

The observed distribution is clearly different from the hypothetical distribution with zero true correlation. Both high positive and high negative correlations are much more frequent in the actual distribution than in the reference distribution. (If we combine the three upper classes, .71 - 1.00 and also combine the three lover classes, -.71 -- 1.00, in order to obtain classes with hypothetical frequencies no less than 5, we may compare the two distributions by a regular χ^2-test. The χ^2-statistic is 2 646,6, and we have 16 classes. In the χ^2-distribution only 1 per cent of the observations will have a value of above 30.6 when there are 15 degrees of freedom.) We can thus conclude that there is evidence of both substitution and complementarity in our figures.

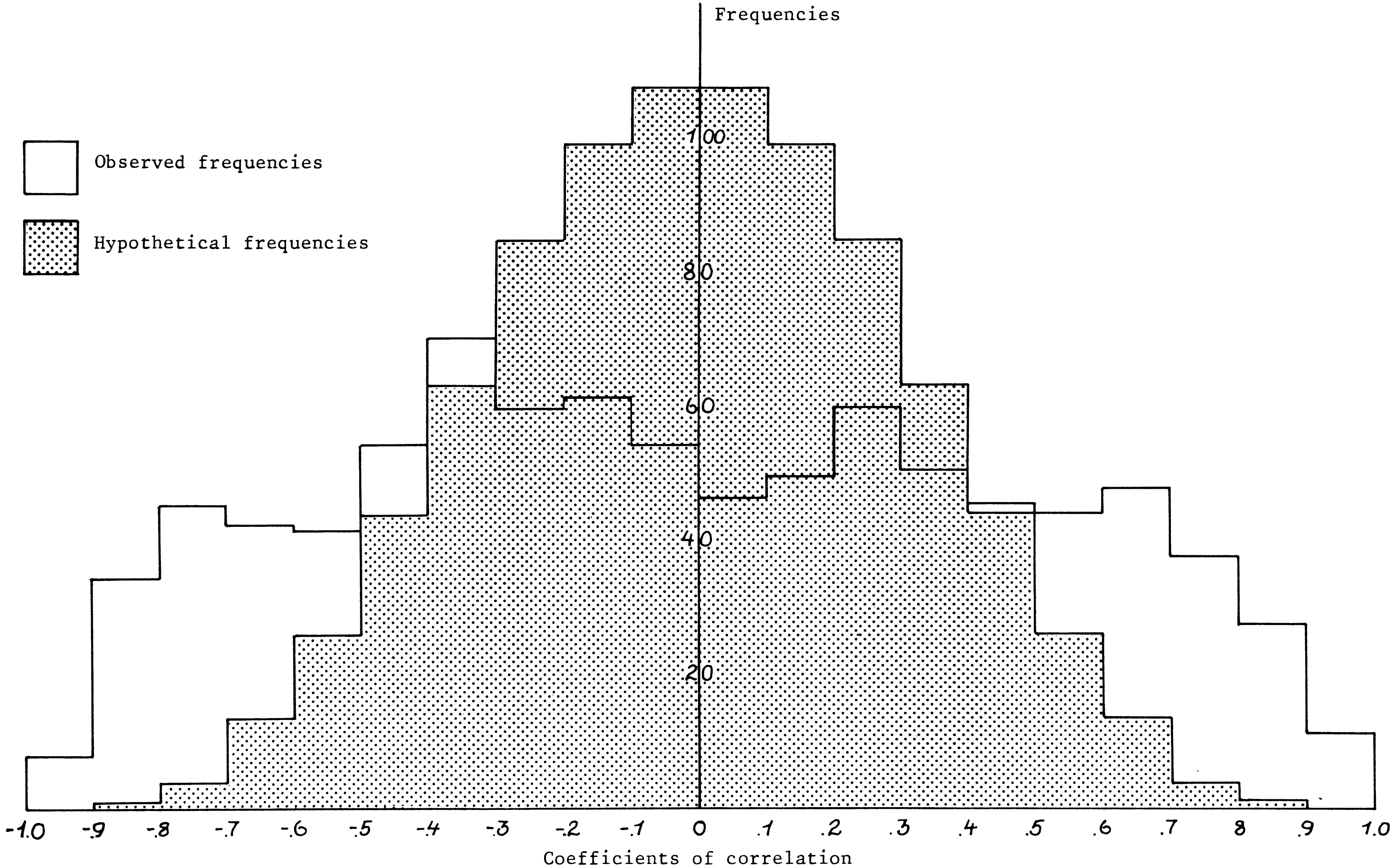

Figure 1. Coefficients of correlation between 883 pairs of input-output ratios.

Table 3. Coefficients of correlation between 883 pairs of input-output ratios

Size of soefficient of correlation	Observed frequencies	Hypothetical frequencies under 0-correlation
.91 - 1.00	11	-
.81 - .90	27	1
.71 - .80	37	4
.61 - .70	47	12
.51 - .60	44	26
.41 - .50	45	44
.31 - .40	50	63
.21 - .30	59	85
.11 - .20	49	99
(+) 0 - .10	46	107½
(-) 0 - - .10	54	107½
-.11 - - .20	61	99
-.21 - - .30	59	85
-.31 - - .40	70	63
-.41 - - .50	54	44
-.51 - - .60	41	26
-.61 - - .70	42	12
-.71 - - .80	45	4
-.81 - - .90	34	1
-.91 - - 1.00	8	-
Total	883	883
.57 - 1.00	140 (15.9%)	22 (2½ %)
-.56 - .56	592 (67.0%)	839 (95%)
-.57 - - 1.00	151 (17.1%)	22 (2½ %)
Total	883	883
(+) 0 - 1.00	415	441½
(-) 0 - - 1.00	468	441½

In our figures there is a slight overweight of negative correlation coefficients, 468 against 415 positive, or 151 in the range -.57 to - 1.00 against 140 in the range .57 to 1.00.

A χ^2 test indicates that we will get an unequality in the distribution on positive and negative items of this magnitude or greater in somewhere between 5 and 10 per cent of the cases if the probabilities for positive and negative items are equal. (If we test the hypothesis that the probability for a positive correlation coefficient ≥ 0.57 is equal to the probability of a negative coefficient ≤ -0.57 each being equal to the average $\frac{140 + 151}{883}$, or if we test the hypothesis that the probability for a positive sign is equal to the probability for a negative sign for each of the 191 correlation coefficients with numerical values above 0.56, we get probabilities of more than $\frac{1}{2}$ for discrepancies at least as great as the one between 140 and 151.)

We have chosen to make a distinction between coefficients with numerical values below 0.57 and those equal to or above this value, because correlation coefficients equal to $\pm$ 0.57 correspond to the critical values at the 5 per cent level in the hypothetical t-distribution with zero true correlation and 10 (= 12-2) degrees of freedom. (The corresponding values at the 1 per cent probability level are $\pm$ 0.71.)

We will now investigate if there is a tendency for the high correlation coefficients to cluster, i.e. if a small number of sectors have relatively many high correlation coefficients, whereas the rest have relatively few. This would imply that only a limited number of sectors are affected by changes of the substitution and complementarity type in their input-output ratios.

Since 33 per cent of the total number of correlation coefficients are high (numerical values of 0.57 and above) we may for each group of sectors with the same total number of coefficients compute the expected number of sectors with each number of high correlation coefficients, assuming the probability that one particular coefficient will be high to be 0.33, and independent of the value of other coefficients for the same sector.

Since the number of sectors in each group with the same number of coefficients are small, we have to pool groups in order to perform a test.

We may consider, then the groups of sectors with 1 to 5 computed correlation coefficients, those with 6 to 10 and those with from 11 to the maximum 28. Within these groups we have the following percentages of high correlation coefficients:

1 to 5 correlation coefficients: 35.2 per cent high (32.7)

6 to 10 correlation coefficients: 34.0 per cent high (36.5)

11 to 28 (more than 10) correlation coefficients: 31.6 per cent high (31.3)

The figures in parantheses are the percentages that we would have obtained for the computed correlation coefficients if we had grouped the sectors according to the numbers of correlation coefficients that <u>might</u> have been computed, had not the number of computations been restricted by the computer program.

The occurence of numerically high correlation coefficients appears to be about the same, irrespective of the number of input-output ratios in a sector.

Omitting sectors with less than three correlation coefficients we obtain the following results, after some grouping:

Table 4. Realised and hypothetical distributions of sectors[1)] according to the frequencies of high correlation coefficients:

Numbers of correlation coefficients n.		Frequencies of high ($\|r\| \geq 0.57$) correlation coefficients 0	1	2	$3-\frac{1}{2}n$	more than $\frac{1}{2}n$	Total
3-6	Realised	12	19	10		9	50
	Hypothetical	11	19	11		9	50
7-28	Realised	2	6	8	25	7	48
	Hypothetical	1	3	7	33	4	48

1) Coefficient groups.

The occurence of high correlation coefficients among sectors with 3 to 6 computed correlations corresponds extremely well to the distribution we would expect if the probability for a high correlation coefficient was 0.33, and independent of the values of other correlation coefficients in the same sector.

For sectors with 7 to 28 computed correlations there is a difference between the realised and the hypothetical distributions; sectors with very few and with very many high coefficients occur more frequently than one should expect. (If we consider as extreme the cases, where less than two or more than half the coefficients are high, the realised distribution between extreme and not-extreme cases is significantly different from the hypothetical distribution at the 1 per cent significance level by a χ^2 test). 16.7 per cent of the sectors with more than 6 computed correlations had less than two high coefficients against an expected percentage of only 8.3, if there had been no tendency to clustering. Correspondingly 14.6 per cent of these sectors had more than half of their correlation coefficients classified as high, against an expected 8.3 per cent. Unfortunately, we cannot conclude from this that there is a tendency for either a large or a quite small proportion of input-output ratios in a sector to be involved in substitution and complementarity changes. The reason is that a model of independence does not make much sense as a reference model here: If the correlation coefficients between one input-output ratio and each of two others are high, the correlation coefficients between the latter two will also tend to be high.

This is easily illustrated by an example in which we assume correlation coefficients to be either ± 1 or 0, and where we assume that the probability for a random correlation coefficient to be ± 1 is 1/3.

In the case of 3 inputs, we will then get the following probabilities:

Table 5 a. Hypothetical and actual frequencies of high correlations. 3 input-output ratios

Number of high correlations	In a binomial distribution		In a distribution of the type discussed in the example		The actual distribution	
0	.30	(8)	.30	(8)	.38	(10)
1	.44	(11)	.59	(15)	.46	(12)
2	.22	(6)	-	-	.08	(2)
3	.04	(1)	.11	(3)	.08	(2)
Total	1.00	(26)	1.00	(26)	1.00	(26)
Average frequency of high correlations	1/3		0.309		1/3	

The figures in parantheses give the expected distributions when the number of items is 26 as in the actual distribution of sectors with three inputs.

In the case of 4 inputs, i.e. 6 correlations, we get

Table 5 b. Hypothetical and actual frequencies of high correlations. 4 input-output ratios

Number of high correlations	In a binomial distribution		In a distribution of the type discussed in the example		The actual distribution	
0	.088	(2)	.088	(2)	.118	(2)
1	.263	(4)	.472	(8)	.294	(5)
2	.329	(6)	.148	(3)	.236	(4)
3	.219	(4)	.255	(4)	.176	(3)
4	.083	(1)	-	(-)	.176	(3)
5	.016	(-)	-	(-)	-	(-)
6	.001	(-)	.037	(-)	-	(-)
Total	1.000	(17)	1.000	(17)	1.000	(17)
Average frequency of high correlations	1/3		0.292		1/3	

17 is the number of items in the actual distribution

In the case of 5 inputs, i.e. 10 correlations we get

Table 5 c. Hypothetical and actual frequencies of high correlations. 5 input-output ratios

Number of high correlations	In a binomial distribution		In a distribution of the type discussed in the example		The actual distribution	
0	.018	(-)	.017	(-)	-	(-)
1	.090	(1)	.210	(3)	-	(-)
2	.199	(3)	.264	(4)	.154	(2)
3	.262	(3)	.248	(3)	.461	(6)
4	.225	(3)	.143	(2)	.077	(1)
5	.133	(2)	-	(-)	.231	(3)
6	.055	(1)	.106	(1)	.077	(1)
7	.015	(-)	-	(-)	-	(-)
8	.003	(-)	-	(-)	-	(-)
9	-	(-)	-	(-)	-	(-)
10	-	(-)	.012	(-)	-	(-)
Total	1.000	(13)	1.000	(13)	1.000	(13)
Average frequency of high correlations	1/3		0.281		0.361	

13 is the number of items in the actual distribution.

Obviously, we will not expect our actual observations to correspond to the example, since we do not have correlation coefficients of only the values 0 and 1. However, judging from the three examples given here, the actual distributions appear to be closer to the binomial distributions than to the distributions of the example. It seems to be a fair conjecture that the observed distributions indicate no tendency to clustering beyond that which follows from the mechanism that makes correlation high between two input-output ratios which both are highly correlated with a third.

Our data may give some indications about the extent of "Input substitution", involving only 2 or possibly three input items. We may study this problem by considering the number of delivering sectors involved in substitution and complementarity relationships for each receiving sector. We get the following picture:

Table 6. Distribution of sector (groups)[1] according to the number of input-output ratios related by high correlation coefficients

Number of input ratios in each group	Number of sectors (groups)[1]					
	Total	No high correlations between input-output ratios	High negative correlations between only two input-output ratios	High correlations between only three input-output ratios not all positively correlated	High positive correlations between only two input-output ratios	Other
2	13	8	3	.	2	.
3	25	9	5	4	6	1
4	23	3	5	6	2	7[2]
5	19	2	1	5	2	9[3]
6	15	-	3	3	1	8[4]
7-9	15	-	-	-	-	15[5]
Total	110[1]	22	17	18	13	40

1) The units are here the groups of input-output ratios for a sector, for which correlations have been computed. 2) Of these one with high negative correlation between two input-output ratios and high positive correlation between the remaining two. 3) Of these one with high negative correlation between two input-output ratios and high correlation coefficients involving the remaining three. 4) Of these one with high correlation coefficients between two separate pairs of input-output ratios. 5) Of these one with high negative correlation between two separate pairs of input-output ratios, one with high positive correlation between two input-output ratios and high correlation coefficients involving a group of three others, and one with high positive correlations between two separate pairs of input-output ratios and high negative correlation in a third pair.

Among the 110 groups there are only 17 cases of simple substitution of one input for one other i.e. 15.5 per cent of the total, and among these 17 cases 13 occur in groups where there are only 4 or less specified input-output ratios and none occur in the groups with more than 6 specified inputs.

If we consider also cases where up to three input-output ratios are involved in the substitution, we get 35 cases or 31.8 per cent of the total, but there is only 6 or 20 per cent of the 30 groups of more than 5 input-output ratios.

The tabulation confirms an impression that, when the number of input-output ratios in the groups increase, the number of high correlation coefficients also tend to increase, and even when we take into account the cases, referred in the footnote to table 6, where several separate groups of two or three input-output ratios have high correlation coefficients only with input-output ratios in the same group, and thus form separate "association complexes", there does not appear to be any tendency for substitution and complementarity to involve only two or three input-output ratios except in the cases where there are only a small number of inputs alltogether. The existence of what we previously termed "input substitution" does not seem to penetrate our results.

We will then investigate if the substitutions and complementary changes in input-output ratios are smooth and gradual, indicating technological change and spesification change, or if they are more random in relation to time.

From table 7 it will be seen that for 242 or 83.2 per cent of the 291 high correlation coefficients both the correlated input-output ratios were also significantly correlated to time at the 5 per cent level (numerical value of correlation coefficient above 0.57. For 172 or 59.1 per cent both input-output ratios were significantly correlated to time at the 1 per cent level i.e. with the numerical value of the correlation coefficient above 0.70.)

Table 7. Pairs of input-output ratios with high correlation coefficients distributed according to the correlation between the input-output ratios and time

Numerical values of correlation coefficients between input-output ratios and time	Sign of correlation between input-output ratios	Number og pairs	Per cent of total
Both less than .57	Positive	7	2.4
	Negative	5	1.7
	All	12	4.1
One less than .57 one in the range .57 - .70	Positive	7	2.4
	Negative	5	1.7
	All	12	4.1
One less than .57 one greater than .70	Positive	9	3.1
	Negative	16	5.5
	All	25	8.6
One less than .57 alltogether	Positive	23	7.9
	Negative	26	8.9
	All	49	16.8
Both in the range .57 - .70	Positive	1	0.3
	Negative	6	2.1
	All	7	2.4
One in the range .57 - .70 one greater than .70	Positive	40	13.8
	Negative	23	7.9
	All	63	21.7
Both greater than .70	Positive	76	26.1
	Negative	96	33.0
	All	172	59.1
Both greater than .56 alltogether	Positive	117	40.2
	Negative	125	43.0
	All	242	83.2
All	Positive	140	48.1
	Negative	151	51.9
	All	291	100.0
One greater than .70 alltogether	Positive	125	43.0
	Negative	135	46.4
	All	260	89.4

This may be compared with the distribution for the remaining 592 small fo the 883 computed correlation coefficients. The comparison is made in table 8, where the characterisation of correlation with time is based on the numerical size of the coefficients of regression with respect to time for the input-output ratios.

Table 8. Correlated pairs of input-output ratios distributed according to the size of the correlation coefficients and the sizes of the regression coefficients for each input-output ratio with respect to time

Numerical values of regression coefficients measured by their standard deviations	Large correlation coefficients		Small correlation coefficients		All correlation coefficients	
	Number of pairs	Per cent	Number of pairs	Per cent	Number of pairs	Per cent
Both less than 2.2 (No trends)	13	4.5	136	23.0	149	16.9
One less than 2.2 one in the range 2.2 - 3.2 (Moderate trend in one input-output ratio, no trend in the other)	12	4.1	114	19.2	126	14.3
One less than 2.2 one greater than 3.2 (Clear trend in one input-output ratio, no trend in the other)	28	9.7	278	47.0	306	34.6
One less than 2.2 alltogether (no trend in at least one of the input-output ratios)	53	18.3	528	89.2	581	65.8
Both in the range 2.2 - 3.2 (Moderate trends in both input-output ratios)	10	3.4	8	1.3	18	2.0
One in the range 2.2 - 3.2 one greater than 3.2 (Clear trend in one input-output ratio, moderate trend in the other)	68	23.3	39	6.6	107	12.1
Both greater than 3.2 (Clear trend in both input-output ratios)	160	55.0	17	2.9	177	20.1
Both greater than 2.2 (Moderate or clear trend in both input-output ratios)	238	81.7	64	10.8	302	34.2
Total	291	100.0	592	100.0	883	100.0

(For independent, normally distributed variables, the numerical value of the regression coefficient may be expected to exceed 2.2 times its estimated standard deviation in 5 per cent of all cases, and to exceed 3.2 times this standard deviation in 1 per cent of all cases. The grouping in tables 7 and 8 should be identical, but there are some discrepancies due to rounding.)

We see that among the highly correlated pairs of input-output ratios both of them are also significantly correlated to time in 82 per cent of the cases (at the 5 per cent level, 55 per cent of the cases at the 1 per cent level). The corresponding percentage(s) for the pairs of input output ratios with small correlation coefficients is only 11 (3). This must be taken as a strong indication that the majority of substitution and complementarity effects that we can distinguish in our data are caused by gradual unidirectional changes in coefficients.

As can be seen from table 7, there does not seem to be systematic differences between negatively and positively correlated input-output ratios in respect to correlation between input-output ratios and time.

Alltogether there were 49, or 16,8 per cent of the correlations which were numerically higher than .57 where not both input-output ratios were significantly correlated with time at the 5 per cent level, but for 20 of these, or 6,9 per cent, none of the two correlation coefficients with time was below .47.

It thus appears that the majority of cases of high correlation between input-output ratios are associated with gradual monotonic changes over time in these ratios.

Non-proportionalities between inputs and outputs does not appear to be a significant cause of high correlations between input-output ratios. Among the 291 cases of numerically high correlations, there was only one in which both input items in the pair was characterised as not directly proportional to output in our test of the form of the relationships between inputs and outputs. (A report on this testing is under preparation.) In this pair one of the input-output ratios was not significantly correlated with time at the 5 per cent level.

It must also be fair to conclude that what we have called substitution proper and product mix fluctuations must play an unimportant role as causes of numerically high correlations between input-output ratios in our data.

We will next investigate whether the sizes of the input-output ratios have any influence on the occurence of substitution and complementarity effects.

The results of grouping the correlated pairs of input-output ratios according to the size of the correlation coefficient and according to the sizes in per cent of the input-output ratios are given in tables 9 and 10.

Table 9 A. Correlated pairs of input-output ratios distributed according to the size of the coefficient of correlation and according to the size of the input-output ratios. Absolute frequencies

Size in per cent of one input-output ratio	Size in per cent of other input output ratio	Large correlation coefficients Positive	Large correlation coefficients Negative	Large correlation coefficients All	Small correlation coefficients	Total
0 - 2.00	0 - 2.00	24	17	41	77	118
	2.01- 5.00	29	31	60	135	195
	5.01-10.00	13	13	26	75	101
	10.01-25.00	10	11	21	39	60
	25.01-50.00	2	3	5	15	20
	50.01 and over	1	3	4	7	11
	All	79	78	157	348	505
2.01- 5.00	2.01- 5.00	17	16	33	61	94
	5.01-10.00	15	14	29	64	93
	10.01-25.00	11	14	25	45	70
	25.01-50.00	4	3	7	18	25
	50.00 and over	1	3	4	9	13
	All 1)	77	81	158	332	490
5.01-10.00	5.01-10.00	5	7	12	11	23
	10.01-25.00	5	4	9	24	33
	25.01-50.00	2	5	7	5	12
	50.01 and over	-	1	1	2	3
	All 1)	40	44	84	181	265
10.01-25.00	10.01-25.00	1	1	2	5	7
	25.01-50.00	-	2	2	-	2
	50.01 and over	-	2	2	-	2
	All 1)	27	34	61	113	174
25.01-50.00	25.00-50.00	-	1	1	-	1
	All 1)	8	12	20	38	58
50.01 and over	All 1)	2	9	11	18	29
Total		140	151	291	592	883

1) Note that these sums include some items also included in the sums above.

Table 9 B. Correlated pairs of input-output ratios distributed according to the size of the coefficient of correlation and according to the size of the input-output ratios. Relative frequencies

Size in per cent of one input-output ratio	Size in per cent of other input output ratio	Large correlation coefficients: Positive	Large correlation coefficients: Negative	Large correlation coefficients: All	Small correlation coefficients	Total
0 - 2.00	0 - 2.00	20.3	14.4	34.7	65.3	100.0
	2.01- 5.00	14.9	15.9	30.8	69.2	100.0
	5.01-10.00	12.9	12.9	25.8	74.2	100.0
	10.01-25.00	16.7	18.3	35.0	65.0	100.0
	25.01-50.00	10.0	15.0	25.0	75.0	100.0
	50.01 and over	9.1	27.3	36.4	63.6	100.0
	All	15.6	15.5	31.1	68.9	100.0
2.01- 5.00	2.01- 5.00	18.1	17.0	35.1	64.9	100.0
	5.01-10.00	16.1	15.1	31.2	68.8	100.0
	10.01-25.00	15.7	20.0	35.7	64.3	100.0
	25.01-50.00	16.0	12.0	28.0	72.0	100.0
	50.01 and over	7.7	23.1	30.8	69.2	100.0
	All[1)]	15.7	16.5	32.2	67.8	100.0
5.01-10.00	5.01-10.00	21.7	30.4	52.1	47.9	100.0
	10.01-25.00	15.1	12.1	27.2	72.8	100.0
	25.01-50.00	16.7	41.7	58.4	41.6	100.0
	50.01 and over	-	33.3	33.3	66.7	100.0
	All[1)]	15.1	16.6	31.7	68.3	100.0
10.01-25.00	10.01-25.00	14.3	14.3	28.6	71.4	100.0
	25.01-50.00	-	100.0	100.0	-	100.0
	50.01 and over	-	100.0	100.0	-	100.0
	All[1)]	15.5	19.6	35.1	64.9	100.0
25.01-50.00	25.01-50.00	-	100.0	100.0	-	100.0
	All[1)]	13.8	20.7	34.5	65.5	100.0
50.01 and over All[1)]		6.9	31.1	38.0	62.0	100.0
Total		15.9	17.1	33.0	67.0	100.0

1) Note that these sums include some items also included in the sums above.

Table 10. Percentages of large correlation coefficients between input-output ratios distributed according to the size of the input-output ratios

<table>
<tr><th rowspan="2">Size of one input-output ratio in per cent</th><th rowspan="2"></th><th colspan="5">Size of other input-output ratio in per cent</th></tr>
<tr><th>0-
2.00</th><th>2.01-
5.00</th><th>5.01-
10.00</th><th>10.01-
25.00</th><th>25.01 and
over</th></tr>
<tr><td rowspan="3">0 - 2.00</td><td>Positive</td><td>20.3</td><td>14.9</td><td>12.9</td><td>16.7</td><td>9.7</td></tr>
<tr><td>Negative</td><td>14.4</td><td>15.9</td><td>12.9</td><td>18.3</td><td>19.3</td></tr>
<tr><td>Total</td><td>34.7</td><td>30.8</td><td>25.8</td><td>35.0</td><td>29.0</td></tr>
<tr><td rowspan="3">2.01 - 5.00</td><td>Positive</td><td>.</td><td>18.1</td><td>16.1</td><td>15.7</td><td>13.2</td></tr>
<tr><td>Negative</td><td>.</td><td>17.0</td><td>15.1</td><td>20.0</td><td>15.7</td></tr>
<tr><td>Total</td><td>.</td><td>35.1</td><td>31.2</td><td>35.7</td><td>28.9</td></tr>
<tr><td rowspan="3">5.01 - 10.00</td><td>Positive</td><td>.</td><td>.</td><td>21.7</td><td>15.1</td><td>13.3</td></tr>
<tr><td>Negative</td><td>.</td><td>.</td><td>30.4</td><td>12.1</td><td>40.0</td></tr>
<tr><td>Total</td><td>.</td><td>.</td><td>52.1</td><td>27.2</td><td>53.3</td></tr>
<tr><td rowspan="3">10.01 and over</td><td>Positive</td><td>.</td><td>.</td><td>.</td><td colspan="2">8.3</td></tr>
<tr><td>Negative</td><td>.</td><td>.</td><td>.</td><td colspan="2">50.0</td></tr>
<tr><td>Total</td><td>.</td><td>.</td><td>.</td><td colspan="2">58.3</td></tr>
<tr><td rowspan="3">0 - 5.00</td><td>Positive</td><td colspan="2">17.2</td><td colspan="3">15.4</td></tr>
<tr><td>Negative</td><td colspan="2">15.7</td><td colspan="3">16.3</td></tr>
<tr><td>Total</td><td colspan="2">32.9</td><td colspan="3">30.8</td></tr>
<tr><td rowspan="3">5.01 and over</td><td>Positive</td><td colspan="2">.</td><td colspan="3">15.7</td></tr>
<tr><td>Negative</td><td colspan="2">.</td><td colspan="3">27.7</td></tr>
<tr><td>Total</td><td colspan="2">.</td><td colspan="3">43.4</td></tr>
</table>

From table 10 it looks as if the tendency to complementarity becomes somewhat less when the input-output ratios increase. The tendency to complementarity seems to be relatively low as soon as one of the input-output ratios exceeds 10 per cent, more or less irrespective of the size of the other ratio. The tendency to substitution effects, on the other hand, appears to increase with the size of the input-output ratios, and to be somewhat more depending on the sizes of both the input-output ratios involved.

The net effect is that the tendency to association (numerically large correlation coefficients) is somewhat stronger when both input-output ratios are big than when at least one of them is small. The picture is, however, not uniform, and one should perhaps be careful not to rely too much on these conclusions.

Substitutability or complementarity could either be characteristic of a group of goods in a variety of uses, or be characteristic of one particular group of goods in one particular use. There is also the third possibility that one particular type of goods may be substitutable for or complementary to others in a variety of uses, but related to different other goods in different uses.

The occurence of the various types of association will naturally depend on the level of aggregation both in terms of goods, i.e. inputs, and in terms of uses - that is in our case: the sector specification for the industries using the inputs.

It would be of great interest to find out if there are particular groups of inputs, which are associated one way or the other in a variety of uses.

We may investigate for each sector if its deliveries are strongly associated with inputs from other sectors in the various user sectors more or less often than normal. The results might give some indications about whether the observed tendencies to association are related to the products of specific sectors or more randomly distributed. They would also say something about the relative importance of the sectors in the process of economic growth.

We may also investigate for each pair of delivering sectors if the frequency of high positive or negative correlations between input-output ratios originating from the pair is greater or smaller than normal. This would tell us something about the existence of mutually substitutable or complementary pairs of sectors.

We will look into both these problems, but, unfortunately, our data are not sufficient to lead us a long way towards conclusions. It is not easy, on the basis of our data, to arrive at conclusions regarding the occurence of the various types of association for individual input delivering sectors. We have found that about 1/3 of all the computed coefficients of correlation are high, and that the frequencies of positive and negative covariations are about equal, both when the numerical value of the correlation coefficient is high and when it is low.

However, since we can only meaningfully compute the coefficients of correlation between input-output ratios for deliveries into the same using sector, the majority of conceivable combinations of delivering sectors do not occur at all, or only once or twice each in our data. The basis for evaluating the frequencies of high correlations for particular combinations of delivering sectors is consequently limited.

In Appendix table A we consider each sector (Norwegian or foreign non-competitive, combined Norwegian and foreign competitive) in its function as provider of inputs to other sectors. We examine then the frequency of high negative and high positive coefficients of correlation between pairs of input-output ratios, where the given sector is provider of one of the inputs in each pair.

In order to have standards against which to evaluate the observed frequencies of high coefficients of correlation, we have computed the probabilities of obtaining deviations from the expected values, at least as large as the observed ones, if the probability of obtaining a numerically high value ($\geq$ 0.57) of the coefficient of correlation were o.33 and if the probability of obtaining a high negative value were 0.165 and that of obtaining a high positive value were also 0.165, and the two latter probabilities were independent. In computing the probabilities for high negative correlations we have disregarded the observed frequencies of high positive correlations, and in computing the probabilities for obtaining high positive correlations we have disregarded the observed frequencies of high negative correlations. The probabilities for the deviations from the expected values for numerically high coefficients of correlation have been computed from binomial distributions for sectors with less than 13 observations whereas the normal distribution has been used as an approximation for sectors with 13 or more observations. The probabilities for deviations from the expected values for high negative and for high positive coefficients of correlation have been computed from binomial distributions for sectors with less than 20 obeervations, whereas the normal distribution has been used for sectors with 20 or more observations.

In Appendix table A and in the more aggregated tables 11 a-c, 12 and 13 we have grouped the sectors according to the numerical values of the computed probabilities. In Appendix table A and in table 12 we have counted as high, respectively low, frequencies with probabilities less than 0.30. In tables 11 a-c and B we have counted as very high, respectively very low, frequencies with probabilities of 0.05 and less and as high, respectively low frequencies with probabilities 0.06-0.29.

We leave it to the reader to contemplate the characteristics of individual industries in regard to substitutability and complementarity, on the basis of the grouping in Appendix table A.

The aggregated tables 11 a-c, 12 and 13 demonstrate that there are considerable differences between the sectors in regard to the occurence of high coefficients of correlation, but according to table 13, the variations do not appear to exceed appreciably what we should expect on the basis of a theory of no "clustering".

Of the 62 sectors, for which we have observations (sectors with specified input deliveries to using sectors which have specified input deliveries from at least one other delivering sector), there are 7, or 11.3 per cent with frequencies for numerically high correlations which deviate more from their expected value than one should expect in 5 per cent of all cases. There are 4, or 6,5 per cent with frequencies for high negative correlations deviating as much, and 5, or 8,1 per cent for high positive correlations.

Table 11 a. Frequencies of numerically high correlation coefficients

Frequencies of numerically high correlation coefficients [1]	Number of sectors	Number of observations	High coefficients of correlation Number	Per cent
Very high (deviations of probability 0.05 and less)	3	82	44	53.7
High (deviations of probability 0.06-0.29)	13	385	157	40.8
Normal (deviations of probability 0.30 and more)	36	728	246	33.8
Low (deviations of probability 0.06-0.29)	6	367	95	25.9
Very low (deviations of probability 0.05 and less)	4	204	40	19,6
Total	62	1 766	582	33.0

1) The classification is in terms of deviations from the expected number of numerically high coefficients of correlations if the probability for a high coefficient in each observation were 0.33 and independent of other observations. The deviations are measured in terms of the probabilities of obtaining (positive and negative) deviations of at least the given magnitudes when the individual probabilities are 0.33 and independent.

Table 11 b. Frequencies of high negative correlation coefficient.

Frequencies of high negative correlation coefficients 1)	Number of sectors	Number of observations: Total	Number of observations: High negative	Percentage of high negative coefficients
Very high (deviations of probability 0.05 and less)	2	41	16	39.0
High deviations of probability 0.06-0.29)	8	185	50	27.0
Normal (deviations of probability 0.30 and more)	48	1 382	228	16.5
Low (deviations of probability 0.06-0.29)	2	77	4	5.2
Very low (deviations of probability 0.05 and less)	2	81	4	4.9
Total	62	1 766	302	17.1

1) The classidication is in terms of deviations from the expected number of high negative coefficients of correlation if the probability for a high negative coefficient were 0.165. The deviations are measured in terms of the probabilities of obtaining (positive and negative) deviations of at least the given magnitudes, when the individual probabilities are 0.165 and independent.

Table 11 c. Frequencies of high positive correlation coefficients.

Frequencies of high positive correlation coefficients 1)	Number of sectors	Number of observations: Total	Number of observations: High positive	Percentages of high positive coefficients
Very high (deviations with probability 0.05 and less)	3	90	28	31.1
High (deviations with probability 0.06-0.29)	6	222	54	24.3
Normal (deviations with probability 0.30 or more)	43	863	147	17.1
Low (deviations with probability 0.06-0.29)	8	417	39	9.4
Very low (deviations with probability 0.05 and less)	2	174	12	6.9
Total	62	1 766	280	15.9

1) The classification is in terms of deviations from the expected number of high positive coefficients of correlation if the probability for a high positive coefficient were 0.165 and independent of other observations. The deviations are measured in terms of the probabilities of obtaining (positive and negative) deviations of at least the given magnitudes when the individual probabilities are 0.165 and independent.

Table 12. Frequencies of high negative and high positive correlation coefficients. All sectors

Frequencies 1) of		Number of sectors	Number of observations			Percentages	
high negative correlation coefficients	high positive correlation coefficients		total	high negative	high positive	high negative	high positive
High	High	1	25	7	8	28.0	32.0
High	"Normal"	8	174	52	28	29.9	16.1
"Normal"	High	7	256	37	66	14.4	25.8
"Normal"	"Normal"	34	639	101	113	15.8	16.9
High	Low	1	27	7	2	25.9	7.4
Low	High	1	31	1	8	3.2	25.8
"Normal"	Low	7	525	93	47	17.7	9.0
Low	"Normal"	1	50	3	6	6.0	12.0
Low	Low	2	39	1	2	2.6	5.1
Total		62	1 766	302	280	17.1	15.8

1) Frequencies are considered to be high, respectively low, when the probability of obtaining a frequency deviating at least as much from the expected frequency is < 0.30 in a sample from a population with probability 0.165 for a high negative, respectively positive, coefficient of correlation.

Table 13. Hypothetical and observed distributions of sectors according to the frequencies of high coefficients of correlation

Frequencies of high correlation coefficients [1)]	Hypothetical distribution	Actual distributions		
		for numerically high coefficients	for high negative coefficients	for high positive coefficients
Very high frequencies, probability level 0.05 and less	1.6	3	2	3
High frequencies, probability level 0.06-0.30	7.8	13	8	6
Normal frequencies, probability level above 0.30	43.2	36	48	43
Low frequencies, probability level 0.06-0.30	7.8	6	2	8
Very low frequencies probability level 0.05 and less	1.6	4	2	2
Total	62.0	62	62	62

1) See footnotes to tables 11 a-c. The figures for the "hypothetical distribution" are obtained by dividing the percentages for a given deviation equally between high and low.

We may next ask if there is any connection between the two types of association: Are products which are easily substitutable occuring as complementary to other products more or less often than the average. If we examine the percentages for high negative and high positive correlations in each sector, we get no impression of covariance. This is confirmed if we compute the correlation coefficient between percentages for high negative and for high positive correlation coefficients. This coefficient is only 0.08, or practically zero.

We may now examine those cases where two sectors both occur as deliverers of input to several users, so that we have been able to compute several correlation coefficients between input coefficients for inputs originating from the two sectors. In Appendix tabel B we have listed all sector pairs with more than 4 observations and also all sector pairs with 2 to 4 observations for which no less than 2/3 of the correlation coefficients were either high and negative or high and positive. In the same way as in Appendix table A we have computed percentages and hypothetical probabilities for high correlation coefficients. A more condensed presentation of the same data is given in tables 14-16.

Among the 30 sector pairs with 5 or more observations only two have a frequency of high negative correlations, which correspond to a hypothetical probability of less than 5 per cent, and only one has a corresponding frequency of high positive correlations. If we look at the sector pairs with high frequencies of high negative correlations, some of them can easily be understood to have substitutable products, but more often it is not easy to see right away how the products can be substitutable. Quite often one of the unspecified sectors will be a member of one of these pairs, and one can easily imagine that changes in specification of inputs may give rise to negative correlations between the coefficients for unspecified and those other coefficients, which are directly affected by the extent to which they are reported separately instead of being lumped in unspecified.

It is perhaps more difficult to form an apriori opinion about which complementarities to expect, and one can imagine causes for the actual cases of high frequency of high positive correlations, as well as one could probably do for other pairs. In particular, it is conceivable that the relatively high proportion of pairs where one of the sectors is unspecified even here has to do with changes in specifications: if unspecified <u>as well as</u> certain input types tend to be reported or omitted simultaneously, they will appear as complementary in our data.

Table 14. Frequencies of high correlations between input-output ratios from the same pair of producing sectors

Sector pairs with:	Number of sector pairs	Number of correlation coefficients Total	Of these high Negative	Of these high Positive	Percentages of high correlation coefficients Negative	Positive	Sum
11 - 22 observations	7	102	15	6	14.7	5.9	20.6
6 - 9 observations	11	82	10	12	12.2	14.6	26.8
5 observations	12	60	14	6	23.3	10.0	33.3
4 observations	22	88	14	13	15.9	14.8	30.7
3 observations	45	135	28	28	20.8	20.8	41.5
2 observations	82	164	25	29	15.2	17.7	32.9
1 observation	252	252	45	46	17.9	18.2	36.1
Total	431	883	151	140	17.1	15.9	33.0

Table 15 a. Frequencies of numerically high correlation coefficients between input-output ratios for inputs originating from the same pairs of producing sectors. Sector pairs with 5 or more observations each

Frequencies of numerically high correlation coefficients:	Number of sector pairs	Per cent	Number of observations	High coefficients of correlations Number	Per cent
Very high (Deviations of probability 0.05 and less)	-	-		-	-
High (Deviations of probability 0.06 - 0.29)	-	-	-	-	-
Normal (Deviations of probability 0.30 and more)	24	80.0	168	55	32.7
Low (Deviations of probability 0.06 - 0.29)	4	13.3	34	3	8.8
Very low (Deviations of probability 0.05 and less)	2	6.7	42*	5	11.9
Total	30	100.0	244	63	25.8

Table 15 b. Frequencies of high negative correlation coefficients between input-output ratios for inputs originating from the same pairs of producing sectors. Sector pairs with 5 or more observations each

Frequencies of high negative correlation coefficients	Number of sector pairs	Per cent	Number of observations	High negative coefficients of correlations Number	Per cent
Very high (Deviations of probability 0.05 and less)	2	6.7	10	6	60.0
High (Deviations of probability 0.06 - 0.29)	4	13.3	30	11	36.7
Normal (Deviations of probability 0.30 and more)	23	76.7	184	21	11.4
Low (Deviations of probability 0.06 - 0.29)	1	3.3	20	1	5.0
Very low (Deviations of probability 0.05 and less)	-	-	-	-	-
Total	30	100.0	244	39	16.0

Table 15 c. Frequencies of high positive correlation coefficients between input-output ratios for inputs originating from the same pairs of producing sectors. Sector pairs with 5 or more observations each

	Number of sector pairs	Per cent	Number of observations	High positive coefficients of correlation Number	Per cent
Very high (Deviations of probability 0.05 and less)	1	3.3	5	3	60.0
High (Deviations of probability 0.06 - 0.29)	4	13.3	38	8	21.0
Normal (Deviations of probability 0.30 and more)	23	76.7	168	13	7.7
Low (Deviations of probability 0.06 - 0.29)	2	6.7	33	-	-
Very low (Deviations of probability 0.05 and less)	-	-	-	-	-
Total	30	100.0	244	24	9.8

Table 16. Sector pairs with different frequencies of observations distributed according to the number of high coefficients of correlation. Actual and hypothetical figures

	Number of sector pairs						
		Of these with:					
	Total	No high correlation	One high negative correlation	Two or more high negative correlations	One high positive correlation	Two or more high positive correlations	Both high negative and high positive correlations
Sector pairs with							
11 - 22 observations	7	-	-	2	-	-	5
(Hypothetical[1])		(-)	(0.1)	(0.5)	(0.1)	(0.5)	(5.8)
6 - 9 observations	11	1	2	1	-	3	4
(Hypothetical[1]).....		(0.6)	(1.0)	(1.3)	(1.0)	(1.3)	(5.8)
5 observations	12	1	4	4	1	2	-
(Hypothetical[1])		(1.6)	(2.0)	(1.3)	(2.0)	(1.3)	(3.8)
4 observations	22	3	5	2	6	1	5
(Hypothetical[1])		(5)	(4)	(2)	(4)	(2)	(5)
3 observations	45	8	14	3	6	7	7
(Hypothetical[1])		(13)	(10)	(3)	(10)	(3)	(6)
2 observations	82	35	20	-	20	2	5
(Hypothetical[1])		(37)	(18)	(2)	(18)	(2)	(5)
1 observation	252	162	45	.	45	.	.
(Hypothetical[1])		(169)	(42)	.	(41)	.	.
Total	431	210	90	12	78	15	26
(Hypothetical[1])		(226)	(77)	(10)	(76)	(10)	(32)

1) The figures that would be expected in a trinomial distribution with probability 0.165 for high negative and the same for high positive correlation coefficients.

One could imagine that sector pairs with many observations were producers of input types that were in some ways associated in use. However, this does not appear to be the case. Judging from table 14, it looks as if cases of high correlation coefficients are more frequent among input ratios from sector pairs which occur more seldom as combinations.

(A possible interpretation of this observation might be that the inputs from sector pairs with few observations are frequently inputs of special raw materials for particular production processes and will tend to have correlated movements under "process substitution", whereas the inputs from sector pairs which occur more often will typically be general materials which are used by all processes, and are mainly influenced by random fluctuations. However, our data do not permit a closer investigation of this possibility).

Although there are certainly sector pairs with extreme frequencies of high correlation coefficients, the occurence of such sector pairs in our total population appears to be well within what one might expect if the occurence of high correlations were purely random and independently distributed over the observations of each sector pair. This is illustrated by the tables 15 a-c and 16.

This analysis of substitution and complementarity in relation to the product types involved does not bring out a dramatic distinction between substitutable and non-substitutable, between complementary and non-complementary or even between "associable" and "non-associable" product types. Even though there are differences between sectors, the differences could apparently easily be the result of mere chance.

III. Substitution and complementarity between corresponding Norwegian and imported inputs

To the extent possible we have also tested the correlations between specified input coefficients of the type Norwegian, competitive and the corresponding coefficients for Imports, competitive. Due to the grouping in the computer program, a relatively small number of coefficient pairs have been included in this analysis, namely 34, or 40 per cent of a total of 86 cases where both the input coefficient for Norwegian, competitive and Imports, competitive were large enough to be specified in this study. The distribution of the correlation coefficients is given in table 17 together with the corresponding distribution that would be obtained if these 34 correlation coefficients had been distributed in the same way as the 883 correlation coefficients analysed above.

Table 17. Substitution between Norwegian and corresponding imported inputs

Coefficient of correlation	The observed distribution	Distribution proportional to the distribution of 883 coefficients
0.57 - 1.00	4	5
0(-) - 0.56	8	11
-0.50 - 0(-)	8	11
-0.56 - -0.51	2	1
-1.00 - -0.57	12	6
Total	34	34

As one might expect the proportion of high negative correlations is greater in this group of correlations than in the group of 883 correlations. (The observed distribution, as it is given in table 17, when the classes -0.50 - 0(-) and -0.56 - -0.51 are combined, deviates from a distribution proportional to the distribution of the 883 correlations to an extent so that a deviation of that or greater magnitude could be expected just over 6 times in a hundred, according to the χ^2 distribution). Less than half of these correlations (44 per cent) are significantly negative (i.e. = -0.50 and less) at the 5 per cent level.

As with the 883 correlation coefficients, we may also for these 34 raise the question whether the association effects could be due to or is connected with gradual changes, taking the effects of trends in time: In table 18 we have classified correlation coefficients of above 0.50 as large positive, correlation coefficients of less than -0.50 as large negative and coefficients in between as small. The correlations have then been grouped in accordance with the existence or non-existence of trends in one or both of the correlated input-output ratios, as measured by the value of the regression coefficient of the input-output ratio with respect to time, divided by its standard deviation.

Table 18. Correlation coefficients between input-output ratios for Norwegian and corresponding imported competitive inputs, distributed according to the size of the regression coefficient with respect to time for each input-output ratio[1].

Trend character of input-output ratios	Large positive correlation coefficients		Large negative correlation coefficients		Small correlation coefficients	
	Number of items	Per cent	Number of items	Per cent	Number of items	Per cent
No trends	-	-	4	28.6	5	33.3
Moderate trend in one coefficient, no trend in the other			2	14.3	1	6.7
Clear trend in one coefficient, no trend in the other		-	3	21.4	6	40.0
No trend in at least one of the input coefficients	-		9	64.3	12	80.0
Moderate trend in both coefficients		-		-	-	-
Clear trend in one coefficient, moderate trend in the other	2	40.0	2	14.3	2	13.3
Clear trend in both coefficients	3	60.0	3	21.4	1	6.7
Moderate or clear trend in both coefficients	5	100.0	5	35.7	3	20.0
Total	5	100.0	14	100.0	15	100.0

1) The classification used is:

No trend: regression coefficient less than 2.2 times its standard deviation

Moderate trend: regression coefficient in the range 2.2 - 3.2 times its standard deviation

Clear trend: regression coefficient exceeding 3.2 times its standard deviation

In the five cases of complementarity between Norwegian and corresponding imported inputs the association is connected with trends in the input-output ratios.

Whereas 82 per cent of the large correlations among the 883 observations were between input-output ratios which both had moderate or clear trends, the corresponding percentage of small correlations where both the correlated input-output ratios had moderate or clear trends was 11 among the 883 observations and 20 among the 34 observations here.

We may probably take these results as an indication that the substitution effects which we can observe between Norwegian products and corresponding imports are to a much greater extent the effects of direct shifts between inputs from these two sources than the substitution effects that we can observe between two random input-output ratios.

We can get an impression of the substitutability between the inputs considered here, as compared to substitutability in general, by considering the number of cases in which an input-output ratio for a competitive input has a higher negative correlation with the input-output ratio for the input, with which it is supposed to compete, than with any other input-cutput ratio with which it can be correlated. This number can be compared with what it would have been if each correlation coefficient had the same probability of being the highest. The results are given in table 19.

Table 19. Correlations of input-output ratios with corresponding competitive inputs compared with correlations with other inputs

Number of input-output ratios with which each competitive input-output ratio has been correlated	Number of competitive input-output ratios	Number of cases when correlation with competitive input was negative and higher than all others	Number if all correlations had the same probability of being highest
3	6	2	2.0
4	8	3	2.0
5	4	0	0.8
6	34	13	5.7
7	6	2	0.9
8	10	7	1.2
Total	68	27	12.6

The correlation with the corresponding competitive input is highest in 27 cases, i.e. 39.8 per cent of the total of 68 caese (34 Norwegian, competitive and the corresponding 34 imports, competitive[1]), whereas the expected number would only be about 13 if each correlation had the same chance of being highest. Thus the substitutability between corresponding competitive Norwegian and imported inputs is again confirmed. Still there are 41 cases, or more than 60 per cent, in which at least one other input-output ratio was more strongly negatively correlated with a competitive input-output ratio than the input-output ratio for the corresponding "competing" input. A fair conclusion seems to be, that our data indicate stronger tendencies to substitutability between Norwegian and corresponding imported competitive inputs, than between random pairs of inputs, but not very much stronger.

IV. Covariations of input-output ratios within substitution groups

In our data we classified as "substitution groups" the main input and all inputs which could be expected to be relatively close substitutes for it in a sector. A sector could have several main inputs and thus several substitution groups. We may now ask if there is a stronger tendency to covariation between input-output ratios which belong to the same substitution group than between input-output ratios in general. We have investigated this by studying the distribution of correlation coefficients for input-output ratios belonging to the same substitution groups. If we consider the specified input-output ratios, the results will be influenced by the fact that in most substitution groups both Norwegian competitive and corresponding imports are grouped together, so that any particular tendency to covariance between corresponding competitive inputs may influence the results. As a concequence of this, we have studied the distributions, both when all specified inputs are treated separately, and when Norwegian and corresponding imported inputs have been lumped together.

In a few cases, where the grouping of data in our computer program made it desirable, we have let specified Norwegian or imports, competitive serve as proxies for competitive inputs combined.

The coverage in this part of the study is given by the following table:

---

1) It may be noted that the results for a pair of corresponding competitive inputs are not entirely independent, but it is difficult to appreciate to what extent, if any, this will influence our results.

Table 20. Possible and actually computed correlations for substitution groups

Substitution groups with possible numbers of correlations	Number of substitution groups	Possible number of correlations	Actually computed correlations in groups of 1-5	6-10	11-28	Total	Total in per cent of possible
Specified inputs							
1 - 5	29	55	30	.	.	30	60.0
6 -10	15	106	25	34	.	59	55.7
11 -28	9	154	16	46	36	98	63.6
Total	53	315	71	80	36	187	59.4
With competitive inputs combined							
1 - 5	40	62	54	.	.	54	87.1
6 -10	8	56	14	40	.	54	96.5
Total	48	118	68	40	.	108	91.5

We may now compare the distribution of correlation coefficients for inputs belonging to the same substitution groups with the distribution of all the 883 correlation coefficients studied in the first part of this study and also with the 34 correlation coefficients between Norwegian, competitive and corresponding Imports, competitive. The comparison can be made in table 21.

Table 21. Coefficients of correlation between pairs of input-output ratios

Size of coefficient of correlation	Inputs belonging to the same substitution groups: Specified inputs		Competitive inputs combined		Norwegian and corresponding imported inputs		All specified inputs	
		Pct.		Pct.		Pct.		Pct.
.71 and over	11	5.9	7	6.5	3	8.8	75	8.5
.57 - .70	9	4.8	8	7.4	1	2.9	65	7.4
.31 - .56	34	18.2	13	12.0	4	11.7	121	13.7
.11 - .30	27	14.5	13	12.1	2	5.9	108	12.2
-.10 - .10	24	12.8	12	11.1	3	8.8	100	11.3
-.30 - -.11	15	8.0	13	12.0	3	8.8	120	13.6
-.56 - -.31	41	21.9	23	21.3	6	17.7	143	16.2
-.70 - -.57	15	8.0	6	5.6	6	17.7	64	7.3
-.71 and less	11	5.9	13	12.0	6	17.7	87	9.8
T o t a l	187	100.0	108	100.0	34	100.0	883	100.0
.57 - 1.00	20	10.7	15	13.9	4	11.7	140	15.9
0 - .56	71	38.0	33	30.5	8	23.5	275	31.1
-.56 - 0	70	37.4	41	38.0	10	29.4	317	35.9
-1.00 - -.57	26	13.9	19	17.6	12	35.4	151	17.1

The figures do not appear to confirm our hypothesis of a stronger tendency to covariation for input-output ratios in the substitution groups. Surprisingly, the tendency to covariation appears to be somewhat stronger when we consider the competitive inputs combined than when we consider all the specified inputs in the substitution groups. This is possibly an effect of the aggregation which is done when corresponding competitive inputs are combined.

In any case we must conclude that our efforts at designating particular "substitution groups" around the main inputs into each sectors does not appear to have given us groups of input-output ratios with stronger tendencies to covariation than pairs of input-output ratios picked at random.

Since we found that covariations within pairs of input-output ratios could apparently to a very large extent be explained by the existence of time trends in both ratios, it makes sense to investigate what this factor may mean for covariation in the substitution groups. The existence of linear time trends for pairs of input-output ratios within substitution groups are indicated by table 22. The results for all specified inputs are reproduced from table 8 for comparison.

Table 22. Correlated pairs of input-output ratios within substitution groups, distributed according to the sizes of the regression coefficients with respect to time for each of the input-output ratios[1)]

Trend character of input-output ratios	Input-output ratios in substitution groups												All specified input-output ratios			
	Specified inputs						With competitive inputs combined									
	Large positive coeffi-cients		Large negative coeffi-cients		Small coeffi-cients		Large positive coeffi-cients		Large negative coeffi-cients		Small coeffi-cients		Large coeffi-cients		Small coeffi-cients	
	No.	Pct.	No.	Pct.	No.	Pct.	No.	Pct.	No.	Pct.	No.	Pct.	No.	Pct.	No.	Pct.
No trends	3	14.3	2	8.0	22	15.6	-	-		-	14	18.9	13	4.5	136	23.0
Moderate trend in one coefficient, no trend in the other.	2	9.5	3	12.0	25	17.7	1	6.7	1	5.3	15	20.2	12	4.1	114	19.2
Clear trend in one coefficient, no trend in the other.	3	14.3	3	12.0	63	44.7	1	6.6	1	5.3	33	44.6	28	9.7	278	47.0
No trend in at least one of the input coefficients	8	38.1	8	32.0	110	78.0	2	13.3	2	10.6	62	83.7	53	18.3	528	89.2
Moderate trends in both coefficients..	1	4.8			2	1.4	1	6.7	2	10.5	1	1.4	10	3.4	8	1.3
Clear trend in one coefficient, moderate trend in the other	3	14.3	6	24.0	20	14.2	5	33.3	2	10.5	9	12.2	68	23.3	39	6.6
Clear trend in both coefficients	9	42.8	11	44.0	9	6.4	7	46.7	13	68.4	2	2.7	160	55.0	17	2.9
Moderate or clear trend in both coefficients	13	61.9	17	68.0	31	22.0	13	86.7	17	89.4	12	16.3	238	81.7	64	10.8
T o t a l	21	100.0	25	100.0	141	100.0	15	100.0	19	100.0	74	100.0	291	100.0	592	100.0

1) The classification is: No trend: regression coefficient less than 2.2 times its standard deviation.
Moderate trend: regression coefficient in the range 2.2-3.2 times its standard deviation.
Clear trend: regression coefficient exceeding 3.2 times its standard deviation.

The results for substitution groups correspond very well with the general results. Possibly, the existence of time trends does not explain quite as much of the high correlations when we consider all specified inputs in the substitution groups as when we consider the competitive inputs combined in the substitution groups, or when we consider all specified inputs. The percentages of high correlations characterised as having moderate or clear trend in both input-output ratios, were 65.2, 88.3 and 81.7 in the three groups respectively.

The differences are, however, hardly big enough to confirm the hypothesis which led to our definition of "substitution groups".

V. Covariations of input-output ratios and sums of input-output ratios

A particular aspect of substitution is its effects on the sum of input-output ratios for intermediate goods. Let us term this sum the input sum ratio.

When we single out the intermediate inputs as a separate group and compute the input sum ratio, we may have changes in this sum ratio[1)] as a result of substitutions involving the relative proportions of intermediate inputs and not affecting the input proportions for primary inputs like labour and capital, we may have changes in the input sum ratio, which are associated with changes in relative proportions for both intermediate and primary inputs, and we may have changes in the input sum ratio under constant relative proportions between intermediate inputs with or without changes in the relative use of primary inputs.

If the latter type of changes in the input sum ratio are dominating, this is a support for the method of proportional adjustments of input-output ratios for intermediate goods, on the basis of observed changes in the input sum ratio in cases where observations of individual inputs are lacking.

We have studied this problem by analysing the coefficients of correlation between specified input-output ratios and the input sum ratios. We have selected specified input-output ratios for inclusion in this part of the study in precisely the same manner as we did for the study of substitution and complementarity between inputs in general. We obtain in this way from our computer program altogether 285 correlations, whereas the greatest number which might have been computed is 404, i.e. we cover 70,6 per cent of all. We obtain 193 or 85,7 per cent of the 225 possible correlations between ratios for competitive inputs combined and input sum ratios and an additional 7 cases or 3.1 per cent where we have substituted the ratio for Norwegian,

1) Computed from constant price values.

competitive for competitive inputs combined. We obtain 75 or 49.3 per cent of the 152 correlations between ratios for Norwegian, non competitive and input sum ratios, which could have been computed, and finally 10 or 37.0 per cent of the 27 correlations between ratios for imports, non-competitive and input sum ratios, which could have been computed.

The distribution is given in table 23. There is a considerable overweight of positive correlations. As many as 127 or 44.9 per cent are significantly different from zero at the 5 per cent level, and of these the majority are positive. Since we are here correlating individual addends with their sums, we must expect positive correlations, particularly for items which constitute considerable fractions of the sums with which they are correlated, either because they are big, or because the sum is made up of few addends. In table 24 we have grouped the sectors according to the number of addends in the input sum ratio. For each group we have then ordered the sectors according to the fraction of computed correlations with the input sum ratios which were above 0.50, and given the cumulative distributions, starting from 0 high correlations. The figures indicate a tendency for a greater proportion of the individual input-output ratios to show a strong positive correlation with the input sum ratio, when there are few individual inputs than when there are many, but the differences are perhaps not as big as noe might have expected.

Table 23. Distribution of coefficients of correlation between individual input-output ratios and input sum ratios

Size of coefficient of correlation	Number of coefficients	Per cent
.81 - 1.00	51	17.9
.61 - .80	51	17.9
.41 - .60	36	12.6
.21 - .40	29	10.2
(+)0 - .20	33	11.6
(-)0 - -.20	28	9.8
-.21 - -.40	26	9.1
-.41 - -.60	16	5.6
-.61 - -.80	12	4.2
-.81 - -1.00	3	1.1
T o t a l	285	100.0
.57 - 1.00	108	37.9
-.56 - .56	159	55.8
-.57 - -1.00	18	6.3
T o t a l	285	100.0
(+)0 - 1.00	200	70.2
(-)0 - -1.00	85	29.8
T o t a l	285	100.0

Table 24. Cumulative distributions of sectors according to the proportions of correlation coefficients between individual input-output ratios and the input sum ratio which were high (above 0.50).

Number of addends in the input sum	Fractions of the computed correlations with the input sum ratio which were above .50.										
	0	0-2/10	0-3/10	0-4/10	0-5/10	0-6/10	0-7/10	0-8/10	0-9/10	Total	Number
2 - 3	13.6	13.6	13.6	27.3	50.0	50.0	86.5	86.5	86.5	100.0	22
4	6.7	6.7	26.7	46.7	66.7	66.7	73.4	100.0	100.0	100.0	15
5	18.2	27.3	27.3	63.6	63.6	63.6	81.8	81.8	81.8	100.0	11
6 - 7	7.2	14.3	28.6	57.2	85.7	92.9	92.9	92.9	92.9	100.0	14
8 - 14	14.3	21.5	35.7	71.5	78.6	85.7	92.9	100.0	100.0	100.0	14
T o t a l	11.9	15.8	25.0	50.0	67.2	69.8	85.5	92.1	92.1	100.0	76

The effect of the size of the coefficient is illustrated in table 25, where the correlations have been grouped according to the size of the input-output ratio. The dominance of high, positive coefficients of correlation is evident in all size groups, but, as expected, the big input-output ratios are strongly positively correlated with the input sum ratio considerably more often than the smaller input-output ratios.

Table 25. Correlations between individual input-output ratios and input sum ratios distributed by size of coefficient of correlation and size of input-output ratio

Size of coefficients of correlation	Size of input-output ratio in per cent: Numbers of coefficients: 0- 2.0	2.0- 5.0	5.0- 10.0	10.0- 25.0	25.0- 50.0	50.0 and over	Total	Percentage distributions: 0- 2.0	2.0- 5.0	5.0- 10.0	10.0- 25.0	25.0- 50.0	50.0 and over	Total
.81 - 1.00	4	13	6	16	7	5	51	5.5	12.6	13.6	40.0	43.8	55.6	17.9
.61 - .80	12	18	8	6	5	2	51	16.5	17.5	18.2	15.0	31.2	22.2	17.9
.41 - .60	14	11	7	2	1	1	36	19.2	10.7	15.9	5.0	6.2	11.1	12.6
.21 - .40	7	10	6	4	1	1	29	9.6	9.7	13.6	10.0	6.3	11.1	10.2
0 - .20	8	18	4	3	-	-	33	10.9	17.5	9.1	7.5	-	-	11.6
0 - - .20	5	10	8	5	-	-	28	6.8	9.7	18.2	12.5	-	-	9.8
-.21 - - .40	12	9	3	1	1	-	26	16.4	8.7	6.8	2.5	6.2	-	9.1
-.41 - - .60	7	5	1	2	1	-	16	9.6	4.9	2.3	5.0	6.3	-	5.6
-.61 - - .80	3	7	1	1	-	-	12	4.1	6.8	2.3	2.5	-	-	4.2
-.81 - -1.00	1	2	-	-	-	-	3	1.4	1.9	-	-	-	-	1.1
T o t a l	73	103	44	40	16	9	285	100.0	100.0	100.0	100.0	100.0	100.0	100.0
.57 - 1.00	19	33	15	22	12	7	108	26.0	32.1	34.1	55.0	75.0	77.8	37.9
0 - .56	26	37	16	9	2	2	92	35.6	35.9	36.3	22.5	12.5	22.2	32.3
0 - - .56	23	23	12	7	2	-	67	31.5	22.3	27.2	17.5	12.5	-	23.5
-.57 - -1.00	5	10	1	2	-	-	18	6.9	9.7	2.3	5.0	-	-	6.3
T o t a l	73	103	44	40	16	9	285	100.0	100.0	100.0	100.0	100.0	100.0	100.0

We must conclude that there is every reason to believe that proportional adjustment of the input-output ratios for intermediate goods on the basis of revised estimates of the input sum ratios will improve the estimates of the input-output ratios when data for a past year have to be utilized as a basis and only the input sum ratios can be given for the year of analysis. However, it is a far step from this conslusion to an assertion that the major part of variations in input-output ratios for intermediate goods can be explained by variations in the input sum ratio. As our analysis indicated, the "explanation" appears to be so much depending on the fact that the explanatory" variable is the sum of the "explained" variables, and there is so much "unexplained" variation left in the individual input-output ratios, that one may remain rather sceptical to theories aimed at explaining variations in individual input-output ratios through a theory for the variations in their sum.

Appendix table A. Frequencies of high correlation coefficients. All sectors

Sector	Numbers of correlation coefficients: Total	Of these high: Negative	Of these high: Positive	Percentages of high correlation coefficients: Negative	Positive	Total	Hypothetical probability of obtaining deviations of observed magnitudes from expected numbers for: Negative	Positive	Sum
Sectors with high frequencies for both high negative and high positive correlations:									
1319 Other oil refineries etc.	25	7	8	28.0	32.0	60.0	0.13	0.04	0.004
Sectors with high frequency for high negative and "normal" frequency for high positive correlations:									
1150 Whaling	4	3		75.0	-	75.0	0.01	0.61	0.11
1201 Slaughtering and preparation of meat	6	2	-	33.3	-	33.3	0.27	0.60	1.00
1206 Grain mill products and livestock feed	12	4	3	33.3	25.0	58.3	0.13	0.44	0.07
1251 Sawmills, planing mills and wood preserving	37	13	5	35.1	13.5	48.6	0.003	0.60	0.04
1330 Non-metallic mineral products	42	10	9	23.8	21.4	45.2	0.22	0.40	0.09
1341 Iron and steel works and rolling mills	53	13	8	24.5	15.2	39.7	0.12	0.76	0.31
1740 Railway transport	4	2	1	50.0	25.0	75.0	0.13	1.00	0.11
0055 Unspecified energy	16	5	2	31.3	12.5	43.8	0.17	0.76	0.36
Sectors with normal frequency for high negative and high frequency for high positive correlations:									
1259 Other wood and cork products	27	4	7	14.8	25.9	40.7	0.79	0.20	0.39
1273 Paper, paperboard and cardboard	22	2	7	9.1	31.8	40.9	0.34	0.06	0.43
1318 Vegetable oil mills	34	5	10	14.7	29.4	44.1	0.76	0.05	0.17
1343 Refining of aluminium	20	5	6	25.0	30.0	55.0	0.32	0.11	0.04
1349 Non-ferrous metal foundries	5	-	2	-	40.0	40.0	0.60	0.20	1.00
1380 Building and repairing of steel ships	31	4	10	12.9	32.2	45.1	0.58	0.02	0.15
0052 Unspecified office supplies etc.	117	17	24	14.5	20.6	35.1	0.54	0.26	0.64

Appendix table A (cont.). Frequencies of high correlation coefficients. All sectors

Sector	Numbers of correlation coefficients			Percentages of high correlation coefficients			Hypothetical probability of obtaining deviations of observed magnitudes from expected numbers for		
	Total	Of these high:							
		Negative	Positive	Negative	Positive	Total	Negative	Positive	Sum
Sectors with normal frequencies for both high negative and high positive correlations:									
1130 Hunting etc.	29	6	5	20.6	17.2	37.9	0.53	0.90	0.58
1170 Coal mining	35	7	8	20.0	22.8	42.8	0.60	0.33	0.22
1181 Metal mining	17	4	2	23.5	11.8	35.3	0.50	0.75	0.84
1190 Quarrying and mining, n.e.c.	38	5	8	13.2	21.0	34.2	0.56	0.47	0.87
1202 Dairy products	12	2	2	16.7	16.7	33.3	1.00	1.00	1.00
1203 Margarine	4	1	1	25.0	25.0	50.0	1.00	1.00	0.60
1205 Fish processing	16	3	2	18.8	12.5	31.3	1.00	0.76	0.88
1211 Distilling, rectifying and blending of spirits	9	2	3	22.2	33.3	55.6	1.00	0.37	0.17
1213 Breweries and soft drinks production	7	1	-	14.3	-	14.3	1.00	0.36	0.44
1230 Textiles except knitting and cordage	50	6	6	12.[illegible]	12.0	24.0	0.38	0.38	0.17
1243 Working clothes and other garments	1	-	-	-	-	-	1.00	1.00	1.00
1271 Mechanical pulp	16	2	4	12.5	25.0	37.5	0.76	0.50	0.70
1282 Printing, etc.	12	1	2	8.3	16.7	25.0	0.70	1.00	0.76
1290 Leather and leather products	13	3	1	23.1	7.7	23.1	0.71	0.50	0.87
1300 Rubber products	12	2	3	16.7	25.0	41.7	1.00	0.44	0.69
1311 Calcium carbide and cyanamide	27	4	6	14.8	22.2	37.0	0.79	0.44	0.65
1317 Herring oil and fish-meal.	15	4	4	26.7	26.7	53.4	0.49	0.49	0.09
1342 Iron and steel foundries..	14	2	3	14.3	21.4	35.7	1.00	0.72	0.83
1344 Crude metals not elsewhere classified	74	14	15	18.9	20.3	39.2	0.60	0.41	0.26
1390 Miscellaneous manufacturing	25	3	5	12.0	20.0	32.0	0.53	0.65	0.91
1530 Trade	9	2	3	22.2	33.3	55.6	1.00	0.37	0.17
1552 Non-life insurance	6	-	-	-	-	-	0.60	0.60	0.18
1580 Commercial buildings	34	7	4	20.6	11.8	32.4	0.60	0.44	0.94
1701 Ocean water transport	3	1	1	33.3	33.3	66.7	1.00	1.00	0.26
1702 Coastal water transport ..	7	2	1	28.6	14.3	42.9	0.61	1.00	0.69
1730 Services related to water transport	9	-	1	-	11.1	11.1	0.37	1.00	0.29
1760 Land transport n.e.c.	4	-	-	-	-	-	0.61	0.61	0.31
1780 Services related to transport and storage	7	1	2	14.3	28.6	42.9	1.00	0.61	0.69
1790 Communications	3	-	-	-	-	-	1.00	1.00	0.56
1860 Legal, technical and business services	3	-	1	-	33.3	33.3	1.00	1.00	1.00
1870 Recreation services	4	-	1	-	25.0	25.0	0.61	1.00	1.00
0057 Unspecified services	113	15	17	13.3	17.8	31.1	0.33	0.65	0.29
0033 "Invisible" imports	9	1	2	11.1	22.2	33.3	1.00	1.00	1.00
0094 Transfers	2	-	-	-	-	-	1.00	1.00	0.86

Appendix table A (cont.). Frequencies of high correlation coefficients. All sectors

Sector	Numbers of correlation coefficients: Total	Of these high: Negative	Of these high: Positive	Percentages of high correlation coefficients: Negative	Positive	Total	Hypothetical probability of obtaining deviations of observed magnitudes from expected numbers for: Negative	Positive	Sum
Sectors with high frequency for high negative, and low frequency for high positive correlations:									
1510 Gas supply	27	7	2	25.9	7.4	33.3	0.20	0.20	0.97
Sectors with low frequency for high negative and high frequency for high positive correlations:									
1370 Wires and cables	31	1	8	3.2	25.8	29.0	0.04	0.17	0.64
Sectors with normal frequency for high negative and low frequency for high positive correlations:									
1110 Agriculture	59	12	3	20.3	5.1	25.4	0.45	0.02	0.22
1121 Forestry	24	4	1	16.7	4.1	20.8	1.00	0.10	0.20
1209 Other food preparation ..	36	8	3	22.2	8.4	30.6	0.37	0.18	0.76
1275 Paper and paperboard products	60	12	5	20.0	8.3	28.3	0.49	0.08	0.44
1315 Chemicals and products of chemicals	219	36	26	16.4	11.9	28.3	0.93	0.06	0.14
1340 Ferro alloys	12	2	-	16.7	-	16.7	1.00	0.24	0.36
1356 Metal products except ships	115	19	9	16.5	7.8	24.3	0.97	0.01	0.05
Sectors with low frequency for high negative and normal frequency for high positive correlations:									
1500 Electricity supply	50	3	6	6.0	12.0	18.0	0.04	0.38	0.02
Sectors with low frequencies for both high negative and high positive correlations:									
1140 Fishing etc.	27	1	2	3.7	7.4	11.1	0.07	0.20	0.01
1233 Cordage, rope and twine..	12	-	-	-	-	-	0.24	0.24	0.03
Total	1766	302	280	17.1	15.8	32.8	0.50	0.47	0.40

Appendix table B. Frequencies of high correlations between input ratios from the same pairs of producing sectors

Sector	Numbers of correlation coefficients: Total	Of these high: Negative	Of these high: Positive	Percentages of high correlation coefficients: Negative	Positive	Total	Hypothetical probability of obtaining deviations of observed magnitudes from expected numbers for: Negative	Positive	Sum
1315 Chemicals and products of chemicals .... 1356 Metal products except ships	22	3		13.7	-	13.7	0.59	0.06	0.05
0052 Unspecified office supplies etc. .... 0057 Unspecified services ...	20	1	1	5.0	5.0	10.0	0.25	0.25	0.03
1315 Chemicals and products of chemicals .... 0057 Unspecified services ...	14	1	1	7.1	7.1	14.3	0.49	0.49	0.14
1315 Chemicals and products of chemicals .... 1344 Crude metals n.e.c.	12	4	1	33.3	8.3	41.7	0.13	0.70	0.55
1315 Chemicals and products of chemicals .... 0052 Unspecified office supplies etc.	12	3	1	25.0	8.3	33.3	0.44	0.70	1.00
1110 Agriculture .... 1315 Chemicals and products of chemicals	11	2		18.2	-	18.2	1.00	0.10	0.30
1580 Commercial buildings ... 0052 Unspecified office supplies etc.	11	1	2	9.1	18.2	27.3	0.71	1.00	0.76
1230 Textiles except knitting and cordage .... 1315 Chemicals and products of chemicals	9	-	2	-	22.2	22.2	0.37	1.00	0.73
1580 Commercial buildings ... 0057 Unspecified services ...	9	2		22.2	-	22.2	1.00	0.37	0.73
1315 Chemicals and products of chemicals .... 1330 Non-metallic mineral products	8	-	2	-	25.0	25.0	0.37	0.64	0.73
1315 Chemicals and products of chemicals .... 1500 Electricity supply	8	1		12.5	-	12.5	1.00	0.37	0.29
1344 Crude metals n.e.c. .... 1356 Metal products except ships	8	3	1	37.5	12.5	50.0	0.14	1.00	0.46
1170 Coal mining .... 1315 Chemicals and products of chemicals	7	-				-	0.36	0.36	0.11

Appendix table B (cont.). Frequencies of high correlations between input ratios from the same pairs of producing sectors

Sector	Numbers of correlation coefficients: Total	Of these high: Negative	Of these high: Positive	Percentages of high correlation coefficients: Negative	Percentages of high correlation coefficients: Positive	Percentages of high correlation coefficients: Total	Hypothetical probability of obtaining deviations of observed magnitudes from expected numbers for: Negative	Hypothetical probability …: Positive	Hypothetical probability …: Sum
1315 Chemicals and products of chemicals .... 1341 Iron and steel works and rolling mills	7	1	1	14.3	14.3	28.6	1.00	1.00	1.00
1341 Iron and steel works and rolling mills ...... 1344 Crude metals n.e.c.	7	1	1	14.3	14.3	28.6	1.00	1.00	1.00
1500 Electricity supply ..... 0057 Unspecified services ...	7	-	3	-	42.8	42.8	0.36	0.09	0.69
1275 Paper and paperboard products .... 1315 Chemicals and products of chemicals	6	1		16.7	-	16.7	1.00	0.60	0.67
1341 Iron and steel works and rolling mills ...... 1356 Metal products except ships	6	1	2	16.7	33.3	50.0	1.00	0.27	0.40
1110 Agriculture .... 0052 Unspecified office supplies etc.	5	2		40.0	-	40.0	0.20	0.60	1.00
1121 Forestry .... 1315 Chemicals and products of chemicals	5	2		40.0	-	40.0	0.20	0.60	1.00
1209 Other food preparation.. 1315 Chemicals and products of chemicals	5	1		20.0	-	20.0	1.00	0.60	0.67
1251 Saw mills, planing mills and wood preserving .... 1315 Chemicals and products of chemicals	5	3		60.0	-	60.0	0.04	0.60	0.34
1275 Paper and paperboard products .... 0052 Unspecified office supplies etc.	5	1		20.0	-	20.0	1.00	0.60	0.67
1315 Chemicals and products of chemicals .... 1318 Vegetable oil mills	5	3		60.0	-	60.0	0.04	0.60	0.34
1315 Chemicals and products of chemicals .... 1370 Electrical machinery, apparatus, appliances etc.	5	-	3	-	60.0	60.0	0.60	0.04	0.34

Appendix table B (cont.). Frequencies of high correlations between input ratios from the same pairs of producing sectors

Sector	Numbers of correlation coefficients			Percentages of high correlation coefficients			Hypothetical probability of obtaining deviations of observed magnitudes from expected numbers for		
	Total	Of these high:							
		Negative	Positive	Negative	Positive	Total	Negative	Positive	Sum
1315 Chemicals and products of chemicals .... 1380 Building and repairing of ships	5	-	1	-	20.0	20.0	0.60	1.00	0.67
1344 Crude metals n.e.c. ..... 1370 Electrical machinery, apparatus, appliances, etc.	5	-				-	0.60	0.60	0.18
1356 Metal products except ships .... 0052 Unspecified office supplies etc.	5	1		20.0	-	20.0	1.00	0.60	0.67
1356 Metal products except ships .... 0057 Unspecified services ...	5	1		20.0	-	20.0	1.00	0.60	0.67
0052 Unspecified office supplies etc. .... 0055 Unspecified energy	5	-	2	-	40.0	40.0	0.60	0.20	1.00
S u m, sector pairs with 4 or more correlations	244	39	24	16.0	9.8	25.8	0.82	0.01	0.02
1580 Commercial buildings ... 0055 Unspecified energy	3	3	-	100.0	-	100.0	0.005	1.00	0.04
1319 Other oil refineries etc. 0052 Unspecified office supplies etc.	3	1	2	33.3	66.7	100.0	1.00	0.07	0.04
1580 Commercial buildings ... 0057 Unspecified services ...	2	2	-	100.0	-	100.0	0.03	1.00	0.11
1206 Grain mill products and livestock feed .... 1318 Vegetable oil mills	2	-	2	-	100.0	100.0	1.00	0.03	0.11
1273 Paper, paperboard and cardboard .... 1319 Other oil refineries etc.	2	-	2	-	100.0	100.0	1.00	0.03	0.11
1251 Saw mills, planing mills and wood preserving .... 1259 Other wood and cork products	3	2		66.7	-	66.7	0.07	1.00	0.26
1275 Paper and paperboard products .... 1311 Calcium carbide and cyanamide	3	2		66.7	-	66.7	0.07	1.00	0.26

Appendix table B (cont.). Frequencies of high correlations between input ratios from the same pairs of producing sectors

Sector	Numbers of correlation coefficients			Percentages of high correlation coefficients			Hypothetical probability of obtaining deviations of observed magnitudes from expected numbers for		
	Total	Of these high:							
		Negative	Positive	Negative	Positive	Total	Negative	Positive	Sum
1190 Quarrying and mining n.e.c. .... 1500 Electricity supply	3	-	2	-	66.7	66.7	1.00	0.07	0.26
1230 Textiles except knitting and cordage .... 1390 Miscellaneous manufacturing	3	-	2	-	66.7	66.7	1.00	0.07	0.26
1273 Paper, paperboard and cardboard .... 0052 Unspecified office supplies etc.	3	-	2	-	66.7	66.7	1.00	0.07	0.26
1275 Paper and paperboard products .... 1318 Vegetable oil mills	3	-	2	-	66.7	66.7	1.00	0.07	0.26
1318 Vegetable oil mills .... 0052 Unspecified office supplies etc.	3	-	2	-	66.7	66.7	1.00	0.07	0.26
1319 Other oil refineries etc. 0052 Unspecified office supplies etc.	3	-	2	-	66.7	66.7	1.00	0.07	0.26
1341 Iron and steel works and rolling mills .... 1380 Building and repairing of ships	3	-	2	-	66.7	66.7	1.00	0.07	0.26
1344 Crude metals n.e.c. .... 0052 Unspecified office supplies etc.	3		2	-	66.7	66.7	1.00	0.07	0.26
Sum listed sector pairs with 2 and 3 correlations ...	42	10	22	23.8	52.4	76.2			

TOTAL NUMBER OF SECTOR PAIRS WITH:

4 correlations	22
3 correlations	45
2 correlations	82
1 correlation	252