Shrittier, O.S.

Netherlands School of Economics

ECONOMETRIC INSTITUTE

Report 6601

INFORMATION MEASURES IN THE ANALYSIS OF INTERNATIONAL TRADE

by H. Theil

February 8, 1966 Preliminary and confidential

Chapter 10*

INFORMATION MEASURES IN THE ANALYSIS OF INTERNATIONAL TRADE

10.1. Predicting Trade Flows from Total Exports and Total Imports

The objective of this chapter is the application of information concepts to the analysis of international trade flows, both for prediction and for the measurement of concentration. Prediction will be our first concern.[1] We imagine that the world is divided into n regions and write y_{ij} for the flow of goods and services from the i^{th} region to the j^{th}, measured as a fraction of world trade. Hence the double sum (i, j = 1, ..., n) of the y_{ij} is equal to 1. Furthermore, we write

$$(1.1) \qquad y_{i.} = \sum_{j=1}^{n} y_{ij} \qquad\qquad y_{.j} = \sum_{i=1}^{n} y_{ij}$$

for the total exports of the i^{th} region and the total imports of the j^{th} region, respectively, both expressed as a fraction of world trade.

Suppose that we know total exports and total imports of each region, and hence also $y_{i.}$ and $y_{.j}$, i, j = 1, ..., n. The question is: Can we predict the flow from the i^{th} region to the j^{th}, for all pairs (i, j), from these marginal totals? One very simple method is

$$(1.2) \qquad \hat{y}_{ij} = y_{i.}y_{.j}$$

which amounts to the assumption of import-export independence. It means that the exports from i to j are supposed to be large when i exports much and j imports much, and that the flow from i to j becomes smaller when either i exports little or j imports little or both. Although there is undoubtedly some truth in this assumption, it is clearly of a very approximate nature. Some region i may export little to some other region j in spite of large values of $y_{i.}$ and $y_{.j}$, simply because their distance is large or because of political troubles. However, many of these causes have a more or less permanent character, such as distance and sometimes political troubles. It is therefore conceivable that one can improve on the prediction method by taking account of the distribution of flows in some earlier year.

---

* Chapter 10 of the series of lecture notes on Economic Applications of Information Theory by Henri Theil.

[1] The analysis described in Sections 10.1 and 10.2 is based on P. Uribe, C.G. de Leeuw and H. Theil, "The Information Approach to the Prediction of Interregional Trade Flows," Report 6507 of the Econometric Institute of the Netherlands School of Economics (1965).

Let us assume, therefore, that we have at our disposal numerical data on the flow from i to j in some earlier year, for all pairs (i, j). We write x_{ij} for this flow when measured as a fraction of that year's value of world trade, and

$$(1.3) \qquad x_{i.} = \sum_{j=1}^{n} x_{ij} \qquad x_{.j} = \sum_{i=1}^{n} x_{ij}$$

for the fractions measuring that year's total exports of i and total imports of j, respectively. We shall then usually find that there is no import-export independence in this year. For some pairs (i, j) we have $x_{ij} > x_{i.}x_{.j}$, which means that the exports from i to j are above the independence level, for other pairs $x_{ij} < x_{i.}x_{.j}$ indicating that trade is below the independence level. Equivalently, consider the logarithmic ratio

$$\log \frac{x_{ij}}{x_{i.}x_{.j}}$$

which is the mutual information between the exporting region i and the importing region j (see Section 2.4). This mutual information is positive, zero, or negative when the exports from i to j are above, at, or below the independence level.

When the forces determining the deviations from the independence pattern may be assumed to be approximately constant over time, a rather obvious procedure is to predict on the assumption that the mutual information values do not change from the year of the x's to that of the y's. This leads to the following forecast of y_{ij}, based on the individual flows of the earlier year and the marginal totals $y_{i.}$ and $y_{.j}$:

$$(1.4) \qquad y'_{ij} = \frac{y_{i.}y_{.j}}{x_{i.}x_{.j}} x_{ij}$$

These forecasts should add up to 1 when summed over i and j. Actually, however, they do not satisfy this constraint.[1] But this defect is remedied easily when we divide all forecasts by their sum:

$$(1.5) \qquad y''_{ij} = \frac{y'_{ij}}{\sum_{h=1}^{n} \sum_{k=1}^{n} y'_{hk}}$$

---

[1] Another constraint that may be violated is $y'_{ij} \leq 1$. Take, for example, $x_{ij} = .3$, $x_{i.} = x_{.j} = .4$, $y_{i.} = y_{.j} = .8$. Then $y'_{ij} = 1.2$. Such violations will be rare, however, except when the changes from the year of the x's to that of the y's are very large. It is assumed here that they do not occur.

Even after this adjustment the forecasts are not really satisfactory, because the sum over j of y''_{ij} is in general not equal to the given value $y_{i.}$, nor is the sum over i equal to $y_{.j}$. An additional adjustment is therefore necessary.

The formulation of this second adjustment should be made dependent on the criterion of the quality of the forecasts. A rather obvious criterion is the information inaccuracy:

$$(1.6) \qquad I = \sum_{i=1}^{n} \sum_{j=1}^{n} y_{ij} \log \frac{y_{ij}}{\hat{y}_{ij}}$$

since both the predictions $\hat{y}_{ij}$ and the quantities predicted y_{ij} are nonnegative and add up to 1. Our task is to adjust the forecasts y''_{ij} such that they satisfy the marginal constraints; clearly, we want to change the y''_{ij} as little as possible, because we prefer to retain the underlying idea of constant mutual information values. If our measure for the difference between two sets of probabilities is the information inaccuracy as in (1.6), the obvious adjustment procedure is that of minimizing

$$(1.7) \qquad \sum_{i=1}^{n} \sum_{j=1}^{n} y''_{ij} \log \frac{y''_{ij}}{\hat{y}_{ij}}$$

subject to

$$(1.8) \qquad \begin{aligned} \sum_{j=1}^{n} \hat{y}_{ij} &= y_{i.} \qquad & i = 1, \ldots, n \\ \sum_{i=1}^{n} \hat{y}_{ij} &= y_{.j} \qquad & j = 1, \ldots, n \end{aligned}$$

The $\hat{y}_{ij}$ thus derived may be called <u>two-stage information forecasts</u>. The first stage is based on the assumption of constant mutual information values. This leads to the y'_{ij} of (1.4), followed by the y''_{ij} of (1.5) after the proportional correction. The second stage consists of minimizing, subject to the marginal constraints (1.8), the information inaccuracy (1.7) in which y''_{ij} plays the role of the "observed" value and $\hat{y}_{ij}$ that of the prediction.

To solve the minimization problem we construct the Lagrangian expression

$$\sum_{i=1}^{n} \sum_{j=1}^{n} y''_{ij} \log \frac{y''_{ij}}{\hat{y}_{ij}} - \sum_{i=1}^{n} \lambda_i \Big(\sum_{j=1}^{n} \hat{y}_{ij} - y_{i.} \Big) - \sum_{j=1}^{n} \mu_j \Big(\sum_{i=1}^{n} \hat{y}_{ij} - y_{.j} \Big)$$

We then differentiate with respect to $\hat{y}_{ij}$ (under the assumption that log stands for natural logarithm) and equate the result to zero:

$$-\frac{y''_{ij}}{\hat{y}_{ij}} - \lambda_i - \mu_j = 0 \qquad i, j = 1, \ldots, n$$

The solution of this system of equations (with the unknown $\hat{y}_{ij}$ in the denominator) is rather awkward. The procedure is simplified when we replace the information inaccuracy (1.7) by its leading quadratic term:

$$\text{(1.9)} \qquad \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(\hat{y}_{ij} - y''_{ij})^2}{y''_{ij}}$$

The approximation error is small when the y''_{ij} violate the marginal constraints to a limited extent only (which is usually the case). If we then form a similar Lagrangian expression, differentiate with respect to $\hat{y}_{ij}$, and equate the derivative to zero, we find

$$\text{(1.10)} \qquad \frac{\hat{y}_{ij}}{y''_{ij}} - 1 - \lambda_i - \mu_j = 0 \qquad i, j = 1, \ldots, n$$

which is solved more easily.[1]

10.2. An Application to the Trade of Eight Regions

In the previous section we developed two prediction methods. One is the import-export independence prediction (1.2), which is quite naive and whose performance should therefore be regarded as a lower limit to that of more sophisticated methods. The second is the two-stage information forecast, amended according to the quadratic approximation (1.9). We shall now apply both procedures to import-export data published by J. Waelbroeck for the years 1938, 1948, 1951-52 and 1959-60.[2] These data are based on a division of the world into eight regions; they are presented (in the form of the fractions y_{ij}) in Table 10.1. There are two zeros on the diagonal; they refer to the intra-regional trade of regions which consist of only one country.

The naive prediction method (1.2) requires only data of the prediction year. On comparing (1.2) and (1.6) we see that the information inaccuracy of this method is simply the expected mutual information of the import-export pattern. The values in the four years are as follows:

[1] A straightforward procedure is the Stephan method. See, for example, W.E. Deming, Statistical Adjustment of Data (New York: John Wiley and Sons, Inc., 1944), pp. 121-124.

[2] See J. Waelbroeck, "Une nouvelle méthode d'analyse des matrices d'échanges internationaux," Cahiers économiques de Bruxelles, No. 21 (1964), pp. 93-114. Note that in the analysis described here Japan has been allocated to the Rest of the world, whereas Waelbroeck takes it as a separate region. This is motivated by the zero flow from Other E.E.C. countries to Japan in 1938, followed by positive flows in later years, which would have led to infinite inaccuracy values. Note further that Germany should be interpreted as Western Germany after the War, and that Yugoslavia has been allocated to the Rest of the world, not to the Communist countries.

TABLE 10.1. IMPORT AND EXPORT SHARES IN PERCENTAGE FORM: 1938, 1948, 1951-52, 1959-60

Region of origin	Region of destination								Total exports ($100y_{i.}$)
	1	2	3	4	5	6	7	8	
					1938				
1. North America	3.18	2.19	.91	1.31	3.69	.66	.83	3.98	16.74
2. Latin America	2.27	.45	1.38	.48	1.26	.23	.12	1.11	7.29
3. Germany	.32	1.01	0	2.17	.64	1.50	1.42	2.57	9.62
4. Other E.E.C. countries	.67	.42	1.46	1.48	1.34	.91	.26	2.42	8.97
5. United Kingdom	1.47	.46	.53	.79	0	.90	.53	5.64	10.32
6. Other E.F.T.A. countries	.38	.54	1.16	1.81	1.68	.68	.52	.62	7.42
7. Communist countries	.58	.43	1.62	1.45	1.23	1.23	1.11	2.50	10.15
8. Rest of the world	3.12	.39	3.28	2.75	6.70	.50	2.35	10.40	29.49
Total imports ($100y_{.j}$)	11.98	5.88	10.36	12.24	16.55	6.61	7.14	29.23	100
					1948				
1. North America	6.00	5.79	1.58	3.14	2.33	1.25	.79	6.43	27.31
2. Latin America	4.44	1.05	.21	1.43	1.48	.42	.17	2.16	11.36
3. Germany	.06	.01	0	.64	.17	.28	.02	.06	1.23
4. Other E.E.C. countries	.57	.86	.35	1.98	.90	1.55	.35	3.55	10.11
5. United Kingdom	.98	.85	.27	.83	0	1.16	.21	6.71	10.99
6. Other E.F.T.A. countries	.48	.45	.18	1.35	.87	.86	.50	1.00	5.69
7. Communist countries	.43	.06	.03	.56	.44	.69	2.56	1.66	6.44
8. Rest of the world	3.57	.59	.14	4.76	6.24	1.08	1.67	8.80	26.86
Total imports ($100y_{.j}$)	16.53	9.65	2.76	14.69	12.43	7.29	6.27	30.38	100

					1951-52				
1. North America	6.09	4.60	.88	1.84	1.81	.81	.00	7.51	23.55
2. Latin America	4.46	.79	.41	.87	.74	.35	.04	1.52	9.18
3. Germany	.33	.48	0	1.34	.27	1.08	.08	1.06	4.64
4. Other E.E.C. countries	.90	.75	.98	2.20	1.24	1.42	.25	4.64	12.38
5. United Kingdom	.95	.54	.15	.81	0	.94	.07	5.42	8.88
6. Other E.F.T.A. countries	.50	.50	.70	1.13	1.08	.76	.37	1.05	6.09
7. Communist countries	.12	.03	.11	.33	.32	.42	4.99	1.91	8.24
8. Rest of the world	3.91	.70	1.32	4.39	5.52	.86	1.91	8.43	27.04
Total imports ($100y_{.j}$)	17.28	8.39	4.54	12.92	10.98	6.65	7.71	31.54	100
					1959-60				
1. North America	5.69	3.01	.87	1.84	1.67	.74	.15	6.16	20.12
2. Latin America	3.12	.58	.69	.55	.60	.21	.17	1.04	6.96
3. Germany	.85	.60	0	2.50	.37	1.67	.40	2.37	8.77
4. Other E.E.C. countries	1.30	.67	2.26	2.84	.95	1.95	.91	3.06	13.94
5. United Kingdom	1.29	.38	.34	.79	0	.84	.23	4.10	7.97
6. Other E.F.T.A. countries	.61	.34	1.16	1.07	.91	.91	.38	1.07	6.45
7. Communist countries	.08	.15	.35	.48	.35	.40	7.63	2.62	12.05
8. Rest of the world	3.73	.62	2.02	3.06	4.06	.71	2.06	7.48	23.74
Total imports ($100y_{.j}$)	16.67	6.35	7.68	13.14	8.91	7.43	11.93	27.89	100

1938	.2061 bit
1948	.2540 bit
1951-52	.3479 bit
1959-60	.3532 bit

The figures increase monotonically, which implies a trend away from independence.

As stated above, one should expect smaller inaccuracy values when a two-stage information forecast is used. This is pursued in Table 10.2, which contains these values for each year with each of the earlier years as a base. The figures are all less than the independence values shown above; in the case of 1959-60 with 1951-52 as base the reduction is even larger than 90 per cent.

TABLE 10.2. INFORMATION INACCURACY VALUES (IN BITS) OF TWO-STAGE INFORMATION FORECASTS

Base year	Year to be predicted 1948	1951-52	1959-60
1938	.0995	.1553	.1701
1948		.0906	.1479
1951-52			.0338

The inaccuracy values which we discussed until now refer to the prediction of all individual shares y_{ij} of world trade. We may also be interested in the destination distribution of the exports from a given region i, $y_{ij}/y_{i.}$, $j = 1, \ldots, n$, or in the origin distribution of the imports by a given region j, $y_{ij}/y_{.j}$, $i = 1, \ldots, n$. The corresponding inaccuracy values are

$$I_{i.} = \sum_{j=1}^{n} \frac{y_{ij}}{y_{i.}} \log \frac{y_{ij}/y_{i.}}{\hat{y}_{ij}/\hat{y}_{i.}}$$

(2.1)

$$I_{.j} = \sum_{i=1}^{n} \frac{y_{ij}}{y_{.j}} \log \frac{y_{ij}/y_{.j}}{\hat{y}_{ij}/\hat{y}_{.j}}$$

We have $\hat{y}_{ij}/\hat{y}_{i.} = y_{.j}$ and $\hat{y}_{ij}/\hat{y}_{.j} = y_{i.}$ in the case of the independence prediction (1.2), so that the logarithms in the two formulas of (2.1) are both the logarithm of $y_{ij}/y_{i.}y_{.j}$. Hence $I_{i.}$ is then the average of the mutual information values involving region i as exporter, and $I_{.j}$ is the average of those values which involve the j^{th} region as importer. In the case of the two-stage information forecast the two logarithms in (2.1) can also be simplified; we can write them as $\log (y_{ij}/\hat{y}_{ij})$ because $\hat{y}_{i.} = y_{i.}$, $\hat{y}_{.j} = y_{.j}$.

TABLE 10.3. INFORMATION INACCURACY VALUES (IN BITS) FOR EXPORT SHARE AND IMPORT SHARE PREDICTIONS

Prediction method	North America	Latin America	Germany	Other E.E.C. countries	United Kingdom	Other E.F.T.A. countries	Communist countries	Rest of the World
				Export shares of 1948				
(1.2)	.1567	.2538	1.0533	.1550	.4550	.1757	.8000	.1571
(1.10), base 1938	.0931	.0294	.6754	.0263	.0595	.0960	.4576	.0681
				Export shares of 1951-52				
(1.2)	.2527	.4478	.5006	.1571	.4549	.2250	1.4575	.1123
(1.10), base 1938	.1243	.0442	.2241	.0240	.0752	.0996	.9575	.0627
base 1948	.0972	.0279	.5714	.0238	.0444	.0565	.1295	.0653
				Export shares of 1959-60				
(1.2)	.2814	.4153	.3880	.1751	.3462	.2248	1.1916	.0994
(1.10), base 1938	.1125	.0140	.2365	.0297	.0483	.1501	.7986	.0499
base 1948	.1818	.0215	.5539	.0897	.0772	.0863	.1185	.0957
base 1951-52	.0397	.0449	.0342	.0738	.0296	.0132	.0331	.0092
				Import shares of 1948				
(1.2)	.2403	.4604	.4486	.1058	.3172	.2551	.8121	.1086
(1.10), base 1938	.0210	.0691	1.0126	.0719	.0926	.0480	.4237	.0306
				Import shares of 1951-52				
(1.2)	.3494	.4543	.2597	.1492	.3214	.3425	1.7378	.0834
(1.10), base 1938	.0492	.0817	.1738	.1204	.0970	.1088	1.1267	.0372
base 1948	.0306	.1661	.6079	.0466	.0358	.0249	.2235	.0474
				Import shares of 1959-60				
(1.2)	.3464	.3707	.3775	.1805	.3251	.3892	1.1919	.0684
(1.10), base 1938	.0831	.0882	.1402	.1179	.0767	.1568	.8171	.0301
base 1948	.0736	.2497	.6371	.0886	.0782	.0495	.0942	.1338
base 1951-52	.0122	.0401	.0366	.0149	.0135	.0249	.1146	.0276

The information inaccuracies $I_{i.}$ for the export shares and $I_{.j}$ for the import shares are given in Table 10.3. In general terms they confirm the overall picture of Table 10.2. For example, the two-stage information forecasts are better than the corresponding independence forecasts with only 6 exceptions out of 96. These exceptions all refer to the case in which the early postwar year 1948 is either predicted or used as a base, and most of them deal with Germany between its defeat and its economic recovery. The destination and origin forecasts for 1951-52 are generally better when 1948 is used as a base than when 1938 is used instead. Disregarding Germany, we find that this applies to 11 cases out of 14. Similarly, the forecasts for 1959-60 are better when the 1948 base is replaced by 1951-52; there are only two exceptions to this rule.

We conclude this section with three remarks:

(1) The two-stage information prediction procedure is rather close to the so-called RAS method, which was developed by Stone and Brown for the adjustment of input coefficient matrices in input-output analysis. This relationship is considered in more detail in the Appendix of this chapter (Section 10.A).

(2) The development over time of the individual mutual information values

$$\log \frac{y_{ij}}{y_{i.}y_{.j}}$$

is in many cases also instructive. Take, for example, the Communist countries (region 7), which have substantial $I_{7.}$ and $I_{.7}$ values according to Table 10.3. If we substitute $i = j = 7$ in the mutual information formula, we obtain the following figures:

1938	.619 bit
1948	2.663 bits
1951-52	2.973 bits
1959-60	2.408 bits

These values are all positive, which indicates the obvious fact that the intra-Communist trade is above the independence level. Moreover, the postwar figures are substantially larger than that of 1938, thus indicating that there was an additional concentration of trade within the group. These features are not unknown, of course, but it is interesting to see how they can be measured quantitatively. We shall make use of such information values on a large scale in the last section of this chapter, where the development of the European Economic Community is considered.

(3) The two-stage information prediction procedure has a much wider range of application than foreign trade alone. Take an arbitrary array of nonnegative fractions, x_{ijk}, the triple sum of which is 1. Suppose

that the corresponding values y_{ijk} in some later period is to be predicted, given the marginal values $y_{i..}$, $y_{.j.}$, $y_{..k}$. The first stage leads to the preliminary forecast

$$y'_{ijk} = \frac{y_{i..}}{x_{i..}} \frac{y_{.j.}}{x_{.j.}} \frac{y_{..k}}{x_{..k}} x_{ijk}$$

after which this is adjusted to y''_{ijk} such that the triple sum is 1. The second stage amounts to the minimization of a χ^2, which leads to equations of the type (1.10) with an additional Lagrangian multiplier ν_k.

10.3. Concentration with Respect to Origin, Destination and Composition for Separate Countries and Commodity Groups

In this and the remaining sections of this chapter we shall be concerned with a three-dimensional problem in international trade. Specifically, the subject is the development in the years 1961-1963 of the export-import flows of 183 commodities among eight countries. The countries are listed in Table 10.4. The X's and Y's before their names are their symbols; the X's will be used when the country is the exporter, the Y's when the country is the importer. Since we consider only the trade among these countries, the data to be used cover only part of their total exports and total imports. The last three columns of the table contain the percentage of total exports covered. These percentages range from about 30 to 80; they are on the low side for the United Kingdom and the United States, which is not surprising in view of their important economic ties with countries outside the group of eight considered here.

TABLE 10.4. EIGHT COUNTRIES AND THE PERCENTAGE OF TOTAL EXPORTS COVERED

Country	Symbol	Percentage of exports covered		
		1961	1962	1963
Belgium*	X_1, Y_1	69.1	72.4	75.9
Netherlands	X_2, Y_2	58.1	60.7	67.6
Germany	X_3, Y_3	43.7	45.9	49.0
France	X_4, Y_4	45.2	48.1	49.0
Italy	X_5, Y_5	48.3	51.3	51.3
United Kingdom	X_6, Y_6	30.4	32.6	32.9
Canada	X_7, Y_7	78.8	81.0	77.8
U.S.A.	X_8, Y_8	39.7	39.2	39.6

*Belgium-Luxembourg Economic Union (B.L.E.U.)

The 183 commodities are three-digit code groups used by the O.E.C.D. (Organisation for Economic Co-operation and Development). We shall indicate them by Z_k, k = 1, ..., 183. In some of our explorations these three-digit groups will be aggregated to 10 one-digit groups, to be indicated by S_g, g = 1, ..., 10. These groups are listed in Table 10.5, together with the number of three-digit groups in each one-digit group as well as the percentage share of the one-digit groups in the total trade among the eight countries in 1961. To simplify the terminology we shall speak about commodities and commodity sets when discussing the Z_k and the S_g, respectively. For further details we refer to the Appendix to this chapter (Section 10.B).

TABLE 10.5. DESCRIPTION OF TEN COMMODITY SETS

Symbol	Description	Number of commodities	Percentage share in 1961
S_1	Food and live animals	33	11.4
S_2	Beverages and tobacco	4	2.4
S_3	Crude materials (inedible) except fuels	29	12.9
S_4	Mineral fuels, lubricants and related materials	5	5.4
S_5	Animal and vegetable oils and fats	4	.5
S_6	Chemicals	16	7.1
S_7	Manufactured goods classified chiefly by material	50	25.6
S_8	Machinery and transport equipment	18	26.3
S_9	Miscellaneous manufactured articles	18	7.0
S_{10}	Commodities and transactions not classified according to kind	6	1.3

The following notation will be used. We write p_{ijk} for the exports of commodity Z_k from country X_i to country Y_j, measured as a fraction of the total trade among all eight countries in all 183 commodities. Hence the p_{ijk} are all nonnegative and their triple sum (i, j = 1, ..., 8; k = 1, ..., 183) is equal to 1. Summations are indicated by dots, so that

$$p_{ij.} = \sum_{k=1}^{183} p_{ijk} \qquad p_{.j.} = \sum_{i=1}^{8}\sum_{k=1}^{183} p_{ijk} \qquad p_{..k} = \sum_{i=1}^{8}\sum_{j=1}^{8} p_{ijk}$$

are, respectively, the total exports from X_i to Y_j, the total imports of Y_j, and the total trade in Z_k, all measured as fractions of the eight countries' total mutual trade. Aggregation (partial summation) is indicated by a superscript. For example, we shall write

$$p^3_{ijg} = \sum_{k\epsilon S_g} p_{ijk} \qquad p^3_{..g} = \sum_{k\epsilon S_g} p_{..k} = \sum_{i=1}^{8}\sum_{j=1}^{8}\sum_{k\epsilon S_g} p_{ijk}$$

for the exports of commodity set S_g from X_i to Y_j and for the total trade in S_g, respectively (again, measured as fractions of the value of aggregate trade). The superscript 3 serves to indicate that the third subscript (g) has an aggregative character. Thus, in this notation the last column of Table 10.5 specifies the 1961 values of $100p^3_{..g}$, $g = 1, \ldots, 10$.

Our first objective is to measure the concentration of international trade for each individual country. This problem is similar to that of the industrial concentration which we considered in Sections 8.1 through 8.3; in fact, we shall use the same entropy measure. But there is a difference to the extent that the present concentration problem has a two-dimensional nature for each country. Consider any exporting country X_i; then there is concentration with respect to customers (importing countries $Y_1, \ldots, Y_8$) and also with respect to commodities. For example, it may be that a country's exports are heavily concentrated on one commodity but quite diversified with respect to its customers because this commodity is sold everywhere. Similarly, there is concentration with respect to suppliers and commodities in the case of any country's imports, and with respect to suppliers and customers for any particular commodity.

Specifically, consider any exporting country X_i and its concentration with respect to customers. Its total exports are measured by $p_{i..}$, and the proportions going to the various customers are $p_{i1.}/p_{i..}, \ldots, p_{i8.}/p_{i..}$. The entropy of this distribution is

$$(3.1) \qquad H_{X_i}(Y) = \sum_{j=1}^{8} \frac{p_{ij.}}{p_{i..}} \log \frac{p_{i..}}{p_{ij.}}$$

which is our inverse measure of concentration. For the concentration with respect to commodities we consider $p_{i.1}/p_{i..}, \ldots, p_{i.183}/p_{i..}$, which are the commodity shares of X_i's export basket. The inverse concentration measure is the entropy of this distribution:

$$(3.2) \qquad H_{X_i}(Z) = \sum_{k=1}^{183} \frac{p_{i.k}}{p_{i..}} \log \frac{p_{i..}}{p_{i.k}}$$

For importing countries Y_j we can proceed in a completely analogous manner. We have

$$(3.3) \qquad H_{Y_j}(X) = \sum_{i=1}^{8} \frac{p_{ij.}}{p_{.j.}} \log \frac{p_{.j.}}{p_{ij.}}$$

for the entropy of the distribution over suppliers, and

$$(3.4) \qquad H_{Y_j}(Z) = \sum_{k=1}^{183} \frac{p_{.jk}}{p_{.j.}} \log \frac{p_{.j.}}{p_{.jk}}$$

for the entropy of the distribution over commodities.

TABLE 10.6. ENTROPY VALUES FOR INDIVIDUAL COUNTRIES

Year	Belgium	Netherlands	Germany	France	Italy	United Kingdom	Canada	U.S.A.
				Exports by destination: $H_{X_i}(Y)$				
1961	2.416	2.318	2.639	2.496	2.416	2.689	1.397	2.402
1962	2.425	2.329	2.623	2.453	2.393	2.705	1.283	2.386
1963	2.447	2.344	2.604	2.456	2.435	2.721	1.301	2.396
				Imports by origin: $H_{Y_j}(X)$				
1961	2.485	2.361	2.670	2.454	2.377	2.580	1.063	2.168
1962	2.481	2.391	2.656	2.450	2.406	2.605	.997	2.166
1963	2.487	2.378	2.646	2.473	2.410	2.612	.937	2.174
				Exports by commodities: $H_{X_i}(Z)$				
1961	6.122	6.222	5.847	6.330	5.623	5.911	5.289	6.259
1962	6.135	6.198	5.765	6.326	5.587	5.928	5.379	6.215
1963	6.167	6.200	5.744	6.340	5.631	5.972	5.438	6.186
				Imports by commodities: $H_{Y_j}(Z)$				
1961	6.327	6.394	6.526	6.110	6.244	6.397	6.212	6.129
1962	6.280	6.366	6.472	6.163	6.251	6.407	6.124	6.115
1963	6.275	6.349	6.493	6.188	6.214	6.452	6.104	6.130

				Exports by commodity sets: $H_{X_i}(S)$				
1961	2.237	2.786	2.374	2.697	2.622	2.488	2.380	2.809
1962	2.265	2.804	2.329	2.706	2.631	2.559	2.454	2.758
1963	2.341	2.753	2.315	2.729	2.608	2.593	2.481	2.775
				Imports by commodity sets: $H_{Y_j}(S)$				
1961	2.797	2.669	2.736	2.666	2.683	2.839	2.536	2.677
1962	2.775	2.701	2.731	2.627	2.622	2.833	2.530	2.698
1963	2.746	2.710	2.769	2.657	2.574	2.864	2.527	2.699
				Exports by commodities within sets: $H_{X_iS}(Z)$				
1961	3.885	3.437	3.473	3.634	3.000	3.423	2.909	3.450
1962	3.869	3.395	3.437	3.620	2.956	3.369	2.926	3.457
1963	3.826	3.447	3.429	3.611	3.022	3.379	2.957	3.411
				Imports by commodities within sets: $H_{Y_jS}(Z)$				
1961	3.529	3.725	3.790	3.444	3.561	3.558	3.676	3.452
1962	3.504	3.665	3.741	3.535	3.629	3.574	3.593	3.417
1963	3.529	3.639	3.723	3.531	3.639	3.588	3.577	3.431

The entropies (3.1) through (3.4) are shown in Table 10.6. With the exception of Canada all export entropies by destination vary between 2.3 and 2.7 bits; since the maximum value is log 8 = 3 bits, this range corresponds to a lower limit of almost 80 per cent of the maximum and an upper limit of 90 per cent. The Canadian values are much lower, which indicates the relative insignificance of the Continental European market for that country's exports. The Canadian import entropy by origin is even smaller, and it is also declining. For the United Kingdom and the United States, too, the import entropies (by origin) are below the corresponding export entropies (by destination). For these three countries, therefore, the imports exceed the exports with respect to supplier-customer concentration. We have the opposite picture in the case of Belgium, the Netherlands and Germany, while France and Italy do not show a consistent pattern in all three years.

The export and import entropies by commodities show a more uniform picture. The latter exceed the former in most cases. This reflects the tendency of individual countries to specialize in their production for exports but to diversify their demand for imports, which leads to a positive excess of the export concentration over the import concentration.[1] Nevertheless, there are two exceptions: France and the United States. Their export concentration over commodities exceeds the corresponding import concentration.

We obtain additional insight by applying the aggregation of commodities to commodity sets. Consider the entropy defined in equation (3.2):

$$H_{X_i}(Z) = \sum_{g=1}^{10} \frac{p^3_{i.g}}{p_{i..}} \sum_{k\in S_g} \frac{p_{i.k}}{p^3_{i.g}} \left(\log \frac{p^3_{i.g}}{p_{i.k}} + \log \frac{p_{i..}}{p^3_{i.g}}\right)$$

$$= \sum_{g=1}^{10} \frac{p^3_{i.g}}{p_{i..}} \sum_{k\in S_g} \frac{p_{i.k}}{p^3_{i.g}} \log \frac{p^3_{i.g}}{p_{i.k}} + \sum_{g=1}^{10} \frac{p^3_{i.g}}{p_{i..}} \log \frac{p_{i..}}{p^3_{i.g}}$$

This can be written as

(3.5) $$H_{X_i}(Z) = H_{X_i}(S) + H_{X_iS}(Z)$$

where $H_{X_i}(S)$ is X_i's export entropy by commodity sets:

(3.6) $$H_{X_i}(S) = \sum_{g=1}^{10} \frac{p^3_{i.g}}{p_{i..}} \log \frac{p_{i..}}{p^3_{i.g}}$$

while $H_{X_iS}(Z)$, to be written as

(3.7) $$H_{X_iS}(Z) = \sum_{g=1}^{10} \frac{p^3_{i.g}}{p_{i..}} H_{X_iS_g}(Z)$$

[1] See, for example, M. Michaely, "Concentration of Exports and Imports: An International Comparison," The Economic Journal, Vol. 68 (1958), pp. 722-736.

is the total entropy of X_i's exports over commodities within sets. It is a weighted average of individual within-set entropies:

$$(3.8) \qquad H_{X_i S_g}(Z) = \sum_{k \in S_g} \frac{p_{i.k}}{p^3_{i.g}} \log \frac{p^3_{i.g}}{p_{i.k}}$$

For imports we have an analogous decomposition:

$$(3.9) \qquad H_{Y_j}(Z) = H_{Y_j}(S) + H_{Y_j S}(Z)$$

where

$$H_{Y_j}(S) = \sum_{g=1}^{10} \frac{p^3_{.jg}}{p_{.j.}} \log \frac{p_{.j.}}{p^3_{.jg}}$$

$$(3.10) \qquad H_{Y_j S}(Z) = \sum_{g=1}^{10} \frac{p^3_{.jg}}{p_{.j.}} H_{Y_j S_g}(Z)$$

$$H_{Y_j S_g}(Z) = \sum_{k \in S_g} \frac{p_{.jk}}{p^3_{.jg}} \log \frac{p^3_{.jg}}{p_{.jk}}$$

The right-hand terms of the decompositions (3.5) and (3.9) are shown in the last twelve lines of Table 10.6. It turns out that the export entropies over commodity sets vary between slightly more than 2.2 bits and about 2.8 bits. The maximum value is log 10 = 3.32 bits; hence the lower limit is just below 70 per cent and the upper limit about 85 per cent of the maximum. This percentage range is about the same as that of the export entropies over the individual commodities, for which the maximum is log 183 = 7.52 bits. For the import entropies the percentage ranges are from about 80 to almost 90, which is higher. When considering commodity sets rather than individual commodities, we find that the export concentration exceeds the import concentration in a majority of the cases, but there are several exceptions: the Netherlands, France, United States, and Italy in 1962 and 1963. When considering commodities within sets (see the last six lines of the table), we find that the rule holds except for Belgium, France, and the United States in 1962. We recall in this connection that France and the United States were the only exceptions to this general rule for the total concentration by commodities [the left-hand sides of (3.5) and (3.9)].

There is also the concentration problem of individual commodities with respect to suppliers and customers. The corresponding entropies are

$$H_{Z_k}(X) = \sum_{i=1}^{8} \frac{p_{i.k}}{p_{..k}} \log \frac{p_{..k}}{p_{i.k}}$$

$$(3.11)$$

$$H_{Z_k}(Y) = \sum_{j=1}^{8} \frac{p_{.jk}}{p_{..k}} \log \frac{p_{..k}}{p_{.jk}}$$

Since there are as many as 183 commodities, this would lead to a very large number of entropy values. We shall therefore confine ourselves to a more limited goal by concentrating on commodity sets rather than individual commodities. The entropies are then

$$(3.12)\qquad \begin{aligned} H_{S_g}(X) &= \sum_{i=1}^{8} \frac{p^3_{i.g}}{p^3_{..g}} \log \frac{p^3_{..g}}{p^3_{i.g}} \\ H_{S_g}(Y) &= \sum_{j=1}^{8} \frac{p^3_{.jg}}{p^3_{..g}} \log \frac{p^3_{..g}}{p^3_{.jg}} \end{aligned}$$

which are inverse concentration measures with respect to origin and destination, respectively, of commodity set S_g. These measures are presented in Table 10.7. It turns out that there is more concentration with respect to origin than with respect to destination, which could be expected. The maximum entropy value is log 8 = 3 bits and the largest observed values in the case of destination exceed 2.99 bits (Chemicals, S_6); 11 values (out of 30) exceed 2.9 bits, 20 are 2.8 bits or higher. The origin entropies, on the other hand, are less than 2.8 bits in all cases except five. There is one commodity set (Beverages and tobacco, S_2) for which the origin entropy exceeds the corresponding destination entropy consistently; for one other set (Manufactured goods classified chiefly by material, S_7) the pattern is different in different years. As a whole, however, the picture is very regular.

10.4. Aggregative Measures of Concentration

In the preceding section we measured concentration for a particular exporting country, for a particular importing country, and for a particular commodity set. There exists a natural way to combine such measures. Consider, for example, the composition entropy of X_i's exports as defined in (3.2). If we take the weighted sum of these entropies with the corresponding export shares as weights, we obtain

$$\sum_{i=1}^{8} p_{i..} H_{X_i}(Z) = \sum_{i=1}^{8} p_{i..} \sum_{k=1}^{183} \frac{p_{i.k}}{p_{i..}} \log \frac{p_{i..}}{p_{i.k}} = H_X(Z)$$

$$= \sum_{i=1}^{8} \sum_{k=1}^{183} p_{i.k} \log \frac{1}{p_{i.k}} - \sum_{i=1}^{8} p_{i..} \log \frac{1}{p_{i..}} = H(X, Z) - H(X)$$

In other words, this weighted sum of the composition entropies of individual exports is nothing else than the average conditional entropy of Z given X, that is, $H_X(Z)$, which is equal to the excess of the two-dimensional entropy H(X, Z) over the one-dimensional entropy H(X). This is entirely in accordance with the conditional entropy definition given in Section 3.1. We can regard $H_X(Z)$ as an inverse concentration measure for the exports of all eight countries with respect to their

TABLE 10.7. ENTROPY VALUES FOR TEN COMMODITY SETS

Symbol and description		Entropy of origin: $H_{S_g}(X)$			Entropy of destination: $H_{S_g}(Y)$		
		1961	1962	1963	1961	1962	1963
S_1	Food and live animals	2.561	2.557	2.611	2.766	2.742	2.844
S_2	Beverages and tobacco	2.509	2.591	2.589	2.480	2.544	2.542
S_3	Crude materials (inedible) except fuels	2.398	2.458	2.480	2.817	2.792	2.800
S_4	Mineral fuels, lubricants and related materials	2.621	2.723	2.732	2.927	2.928	2.918
S_5	Animal and vegetable oils and fats	2.186	2.437	2.385	2.794	2.870	2.864
S_6	Chemicals	2.639	2.681	2.709	2.991	2.993	2.997
S_7	Manufactured goods classified chiefly by material	2.877	2.883	2.894	2.876	2.872	2.897
S_8	Machinery and transport equipment	2.591	2.613	2.618	2.933	2.945	2.956
S_9	Miscellaneous manufactured articles	2.774	2.824	2.822	2.897	2.902	2.925
S_{10}	Commodities and transactions not classified according to kind	2.661	2.697	2.661	2.790	2.784	2.757

composition. In the same way we can weight the entropy values (3.4):

$$\sum_{j=1}^{8} p_{.j.} H_{Y_j}(Z) = H_Y(Z) = H(Y, Z) - H(Y)$$

This is an inverse concentration measure of all imports with respect to composition. Consider also:

$$H_{XY}(Z) = H(X, Y, Z) - H(X, Y) = \sum_{i=1}^{8} \sum_{j=1}^{8} \sum_{k=1}^{183} p_{ijk} \log \frac{p_{ij.}}{p_{ijk}}$$

$$= \sum_{i=1}^{8} \sum_{j=1}^{8} p_{ij.} \sum_{k=1}^{183} \frac{p_{ijk}}{p_{ij.}} \log \frac{p_{ij.}}{p_{ijk}} = \sum_{i=1}^{8} \sum_{j=1}^{8} p_{ij.} H_{X_iY_j}(X)$$

This is the weighted average (with weights $p_{ij.}$) of the composition entropies of the individual flows from a particular country to a particular country. We know from Section 3.3 [see equation (3.2) of that section] that $H_{XY}(Z)$ is always smaller than $H_X(Z)$ and $H_Y(Z)$, at least not larger. This is easy to understand in the present context. In the case of $H_{XY}(Z)$ we take both the exporter and the importer as given; it stands to that there is then, on the average, less uncertainty (more concentration) with respect to the composition of the basket.

These derivations show that there are several composition entropies for the trade among the eight countries as a whole, starting with the unconditional entropy H(Z), followed by the single-conditioned entropies $H_X(Z)$ and $H_Y(Z)$, and concluded by the double-conditioned entropy $H_{XY}(Z)$. It is instructive to consider this numerically for one year, 1961 say, for which purpose we present the figures in the following array:

log 183	H(Z)	$H_X(Z)$	$H_Y(Z)$	$H_{XY}(Z)$
7.516	6.673	5.966	6.292	5.559
100	88.8	79.4	83.7	74.0
	100	89.4	94.3	83.3

The first column contains log 183 = 7.516 bits, which is the maximum entropy value in view of the number of commodities. The observed unconditional value H(Z) is 11.2 per cent lower (see the second line, where the maximum is put equal to 100). This entropy measures the dispersion of aggregate trade with respect to composition irrespective of origin and destination. Next is $H_X(Z)$, which measures the dispersion with respect to composition, given the origin but irrespective of the destination. Its value is 10.6 per cent below the unconditional H(Z); see the third line. This difference is a measure for the average gain in knowledge as to the composition when we know the origin of the export baskets. We then have $H_Y(Z)$, which performs the same service except that origin and destination are to be interchanged. The figures show that knowledge of the destination is on the average less useful than knowledge of the origin, which is not really surprising in view of the results obtained in the previous section. Finally, we have $H_{XY}(Z)$

measuring the composition dispersion given both origin and destination. The uncertainty reduction from the level of H(Z) is then 16.7 per cent.

The various unconditional and average conditional entropies with respect to composition are given in the upper half of Table 10.8, both for the composition in terms of the 183 individual commodities and for that in terms of the ten commodity sets. The commodity set entropies are obtained in a completely analogous manner. For example:

$$H_X(S) = H(X, S) - H(X) = \sum_{i=1}^{8} \sum_{g=1}^{10} p^3_{i.g} \log \frac{1}{p^3_{i.g}} - \sum_{i=1}^{8} p_{i..} \log \frac{1}{p_{i..}}$$

$$= \sum_{i=1}^{8} p_{i..} \sum_{g=1}^{10} \frac{p^3_{i.g}}{p_{i..}} \log \frac{p_{i..}}{p^3_{i.g}} = \sum_{i=1}^{8} p_{i..} H_{X_i}(S)$$

see (3.6). The results show that, both with respect to commodities and with respect to commodity sets, knowledge of origin is more informative than knowledge of destination. The unconditional entropy over sets is relatively smaller than that over individual commodities. On the other hand, the additional uncertainty reduction obtained by knowledge of origin or destination or both is comparatively small for sets, so that the double-conditioned entropies $H_{XY}(Z)$ and $H_{XY}(S)$ are both about 25 per cent of the corresponding maximum. These "relative" statements are all based on the last two columns of each array, which contain the entropies in percentage form averaged over the three years.

The lower half of Table 10.8 deals with origin and destination entropies in an analogous manner. The maximum is always log 8 = 3 bits. This is followed by the unconditional entropies, H(X) and H(Y), and by three single-conditioned figures. There are three such figures, not two, because the commodity specification may be either in terms of individual commodities or in terms of commodity sets. In the former case we have

$$H_Z(X) = \sum_{k=1}^{183} p_{..k} \sum_{i=1}^{8} \frac{p_{i.k}}{p_{..k}} \log \frac{p_{..k}}{p_{i.k}}$$

and in the latter:

$$H_S(X) = \sum_{g=1}^{10} p^3_{..g} \sum_{i=1}^{8} \frac{p^3_{i.g}}{p^3_{..g}} \log \frac{p^3_{..g}}{p^3_{i.g}}$$

It stands to reason that $H_Z(X) \leq H_S(X)$, because the Z specification is more detailed than the S specification. In the same way we have $H_{YZ}(X) \leq H_{YS}(X)$, which deal with double-conditioned entropies that are defined in a straightforward manner. The figures indicate that the origin concentration exceeds the destination concentration systematically. In particular, the double-conditioned origin entropy $H_{YZ}(X)$ is very low; it is almost 50 per cent less than the maximum value.

TABLE 10.8. CONDITIONAL AND UNCONDITIONAL ENTROPIES

	1961	1962	1963	Average percentages			1961	1962	1963	Average percentages	
	Commodities						Commodity sets				
log 183	7.516	7.516	7.516	100		log 10	3.322	3.322	3.322	100	
$H(Z)$	6.673	6.631	6.609	88.3	100	$H(S)$	2.767	2.757	2.757	83.1	100
$H_X(Z)$	5.966	5.950	5.963	79.3	89.8	$H_X(S)$	2.567	2.565	2.574	77.3	93.1
$H_Y(Z)$	6.292	6.267	6.273	83.5	94.6	$H_Y(S)$	2.695	2.689	2.694	81.1	97.5
$H_{XY}(Z)$	5.559	5.578	5.609	74.3	84.1	$H_{XY}(S)$	2.465	2.467	2.483	74.4	89.5
	Exports						Imports				
log 8	3	3	3	100		log 8	3	3	3	100	
$H(X)$	2.851	2.875	2.880	95.6	100	$H(Y)$	2.944	2.942	2.960	98.3	100
$H_Y(X)$	2.239	2.245	2.253	74.9	78.3	$H_X(Y)$	2.332	2.312	2.333	77.5	78.9
$H_S(X)$	2.650	2.682	2.698	89.2	93.3	$H_S(Y)$	2.872	2.874	2.897	96.0	97.7
$H_Z(X)$	2.145	2.194	2.235	73.0	76.4	$H_Z(Y)$	2.563	2.578	2.624	86.3	87.8
$H_{YS}(X)$	2.009	2.024	2.041	67.5	70.6	$H_{XS}(Y)$	2.230	2.215	2.241	74.3	75.6
$H_{YZ}(X)$	1.505	1.556	1.589	51.7	54.0	$H_{XZ}(Y)$	1.924	1.941	1.978	64.9	66.1

10.5. Two Country Sets: The Common Market and the Rest

An obvious question is to what extent the import-export data of the eight countries enable us to describe the development of the Common Market in informational terms. It should be stressed at the outset that the analysis which follows gives only a partial answer, since a more complete analysis should include the development of imports and exports during years prior to 1961. The method is therefore more important than the results, although these results are not without interest either.

We shall need some additional symbols indicating aggregation over countries within country sets. There are two such sets: the European Economic Community (E.E.C.) and the Rest. The former will be indicated by Q_1 or R_1, depending on its role as exporter or importer, and it consists of the first five countries: $X_1, \ldots, X_5$ or $Y_1, \ldots, Y_5$. The Rest consists of the last three countries (X_6, X_7, X_8 or Y_6, Y_7, Y_8) and is indicated by Q_2 or R_2. We follow the superscript notation for aggregation over countries also. For example,

$$p^{13}_{rjg} = \sum_{i \in Q_r} \sum_{k \in S_g} p_{ijk} \qquad\qquad p^{12}_{rs.} = \sum_{i \in Q_r} \sum_{j \in R_s} \sum_{k=1}^{183} p_{ijk}$$

are, respectively, the exports of commodity set S_g from the countries of the country set Q_r to country j and the total exports from the countries of country set Q_r to those of country set R_s, both measured as fractions of the aggregate trade among all eight countries. The flows $p^{12}_{rs.}$ are shown in Table 10.9. They indicate that the Common Market share of exports increased from 50.3 to 53.8 per cent, that its import share increased from 54.7 to 58.4 per cent, and that the share of the mutual trade of the E.E.C. countries increased from 36.8 to 41.2 per cent.

TABLE 10.9. SHARES OF TOTAL TRADE OF E.E.C. AND THE REST

Country set of origin	E.E.C.	Rest	Sum	E.E.C.	Rest	Sum	E.E.C.	Rest	Sum
	1961			1962			1963		
E.E.C.	.368	.134	.503	.388	.131	.519	.412	.126	.538
Rest	.179	.319	.497	.174	.307	.481	.172	.290	.462
Sum	.547	.453	1	.563	.437	1	.584	.416	1

We shall now evaluate these developments in informational terms. Consider the following mutual information values:

$$(5.1) \qquad \log \frac{p^{12}_{rs.}}{p^{1}_{r..}\, p^{2}_{.s.}} \qquad\qquad r, s = 1, 2$$

TABLE 10.10. MUTUAL INFORMATION VALUES (5.1) AND (5.2): E.E.C. VERSUS THE REST

Abbreviated description of commodity sets	1961	1962	1963		1961	1962	1963		1961	1962	1963		1961	1962	1963	
	Within E.E.C. (r = s = 1)				From E.E.C. to Rest (r = 1, s = 2)				From Rest to E.E.C. (r = 2, s = 1)				Within Rest (r = s = 2)			
All commodities	422	411	391	−	− 761	− 796	− 831	−	− 606	− 635	− 650	−	500	544	594	+
S_1: Food and live animals	374	389	370		− 779	− 892	− 897	−	− 448	− 480	− 512	−	461	507	559	+
S_2: Beverages and tobacco	518	428	459		− 408	− 410	− 434	−	− 328	− 312	− 361		158	189	206	+
S_3: Crude materials except fuels	750	782	781		−1466	−1517	−1460		− 491	− 591	− 623	−	345	383	391	+
S_4: Mineral fuels, etc.	362	367	314		−1356	−1247	−1151	+	− 828	− 697	− 528	+	950	820	758	−
S_5: Oils and fats	120	122	207	+	− 325	− 316	− 581		− 065	− 096	− 149	−	140	194	276	+
S_6: Chemicals	241	242	235		− 461	− 481	− 504	−	− 285	− 306	− 340	−	344	375	432	+
S_7: Manufactured goods (by material)	450	438	412	−	− 858	− 869	− 908	−	−1034	−1071	−1047		702	739	790	+
S_8: Machinery and transport equipment	369	361	347	−	− 622	− 712	− 770	−	− 499	− 533	− 570	−	434	509	582	+
S_9: Miscellaneous manufactured articles	294	259	227	−	− 363	− 347	− 373		− 789	− 721	− 650	+	501	513	570	+
S_{10}: Other commodities and transactions	460	508	585	+	− 730	− 830	−1038	−	− 562	− 686	− 740	−	424	477	497	+

Note. All figures are to be multiplied by 10^{-3}.

These values, which are completely analogous to the measures used for the intra-Communist trade at the end of Section 10.2 under (2), are shown on the first line of Table 10.10 for each of the three years. They are positive whenever r = s, negative whenever $r \neq s$, which means that the trade within each country set is above the independence level, whereas the trade between the two sets is below that level. For r = s = 2 the successive figures increase monotonically (indicated by + upper right). Hence the trade among the three countries of the Rest increases continuously relative to the independence level. For $r \neq s$ the figures decrease algebraically in the three successive years (indicated by - behind the time series). This means that the trade between the country sets decreases relative to the independence value. It is somewhat surprising to find that the figures for r = s = 1 also decrease, which implies that the trade among the Common Market countries is declining in the direction of the independence level. One should realize, however, that such a development necessarily takes place when the import and export shares of the E.E.C. increase more and more. In the limit we have $p^{12}_{11.} = p^{1}_{1..} = p^{2}_{.1.} = 1$, for which the value of (5.1) for r = s = 1 is zero.[1]

We can also derive similar mutual information values for the separate commodity sets:

$$(5.2) \qquad \log \frac{p^{123}_{rsg}/p^{3}_{..g}}{\dfrac{p^{13}_{r.g}}{p^{3}_{..g}}\;\dfrac{p^{23}_{.sg}}{p^{3}_{..g}}} \qquad r, s = 1, 2 \quad g = 1, \ldots, 10$$

These, too, are given in Table 10.10. We find that they all have the same sign pattern as that of all commodities combined; also, that they increase over time for r = s = 2 with only one exception and that they decrease algebraically for $r \neq s$ in a majority of the cases. [The absence of + or - indicates that the development is not monotonic.] For r = s = 1, dealing with the trade among the Common Market countries, the picture is more diffuse. It is interesting to observe that S_4

[1] The E.E.C. share of the total imports of the E.E.C. increases. It is (.368)/(.547) or 67 per cent in 1961, 69 per cent in 1962, and 71 per cent in 1963. Similarly, the E.E.C. share of the total E.E.C. exports increases: 73, 75, 77 per cent in the three successive years. But when we divide the intra-E.E.C. trade bij the product of E.E.C. imports and exports (all divided by the aggregate value of the trade among the eight countries), the resulting ratio declines over time. In this connection we refer to footnote 1 on page 2, which shows that an increase of the marginal probabilities may lead to a probability of the bivariate distribution which exceeds 1 when the mutual information value is not changed.

(Mineral fuels, lubricants and related materials) is an exceptional commodity set to the extent that its E.E.C. - Rest trade pattern tends to move in the direction of independence.

Another extension of the mutual information formula (5.1) is in the direction of the exports from and the imports by individual countries. When considering total exports from countries to country sets

TABLE 10.11. MUTUAL INFORMATION VALUES (5.3) AND (5.4): E.E.C. VERSUS THE REST FOR INDIVIDUAL COUNTRIES

Country	1961	1962	1963		1961	1962	1963	
	$\log \frac{p^2_{i1.}}{p_{i..}p^2_{.1.}}$				$\log \frac{p^2_{i2.}}{p_{i..}p^2_{.2.}}$			
Belgium	.495	.481	.455	-	-.983	-1.025	-1.061	-
Netherlands	.485	.440	.432	-	-.950	-.884	-.973	
Germany	.405	.393	.375	-	-.716	-.742	-.778	-
France	.442	.446	.417		-.817	-.906	-.919	-
Italy	.246	.272	.246		-.367	-.448	-.437	
United Kingdom	.005	.053	.076	+	-.006	-.071	-.114	-
Canada	-2.359	-2.646	-2.703	-	.980	1.057	1.131	+
U.S.A.	-.352	-.397	-.447	-	.335	.388	.459	+
	$\log \frac{p^1_{1j.}}{p^1_{1..}p_{.j.}}$				$\log \frac{p^1_{2j.}}{p^1_{2..}p_{.j.}}$			
Belgium	.579	.534	.495	-	-.996	-.954	-.932	+
Netherlands	.446	.387	.372	-	-.658	-.582	-.604	
Germany	.394	.413	.393		-.551	-.638	-.653	-
France	.478	.472	.428	-	-.729	-.781	-.741	
Italy	.188	.229	.268	+	-.218	-.296	-.390	-
United Kingdom	-.104	-.104	-.143	-	.098	.105	.150	+
Canada	-2.894	-2.947	-3.069	-	.906	.956	1.018	+
U.S.A.	-.491	-.552	-.587	-	.369	.426	.474	+

we have the following mutual information values:

$$(5.3) \qquad \log \frac{p^2_{is.}}{p_{i..}p^2_{.s.}} \qquad i = 1, \ldots, 8;\ s = 1, 2$$

Similarly, for total imports by countries from country sets:

$$(5.4) \qquad \log \frac{p^1_{rj.}}{p^1_{r..}p_{.j.}} \qquad r = 1, 2;\ j = 1, \ldots, 8$$

These values are given in Table 10.11. It turns out that the E.E.C. countries have a perfectly consistent sign pattern: The information values (5.3) and (5.4) are all positive when the trade is with the E.E.C. group, thus indicating that this trade (both exports and imports) is above the independence level, and the signs are all negative for trade with the Rest group. The three countries of the latter group are divided. For Canada and the United States we have a consistent opposite sign pattern; in particular, the Canadian - E.E.C. mutual information values take substantial negative values (varying from -2.36 to -3.07 bits). For the United Kingdom, on the other hand, the picture is more diffuse. Its exports to E.E.C. and to the Rest were close to the independence level in 1961, but in the later years there was a movement toward Continental Europe and away from the English-speaking countries on the American continent. The imports, on the other hand, drifted toward the latter countries.

Finally, we shall consider a combination of the extensions (5.2) and (5.3)-(5.4) by taking both separate commodity sets and individual countries (either as exporters or as importers). The formulas are:

$$(5.5) \qquad \log \frac{p^{23}_{isg}/p^{3}_{..g}}{\dfrac{p^{3}_{i.g}}{p^{3}_{..g}} \dfrac{p^{23}_{.sg}}{p^{3}_{..g}}} \qquad i = 1, \ldots, 8 \quad s = 1, 2 \quad g = 1, \ldots, 10$$

$$(5.6) \qquad \log \frac{p^{13}_{rjg}/p^{3}_{..g}}{\dfrac{p^{13}_{r.g}}{p^{3}_{..g}} \dfrac{p^{3}_{.jg}}{p^{3}_{..g}}} \qquad r = 1, 2 \quad j = 1, \ldots, 8 \quad g = 1, \ldots, 10$$

Table 10.12 contains a summary; the information values (5.5) are given in the first set of three columns for s = 1 (exports to E.E.C.) and in the second set for s = 2 (exports to the Rest group), the values (5.6) in the third set for r = 1 and in the fourth for r = 2. Our comments are as follows:

(1) The sign pattern is stable in the sense that different signs in different years of the same mutual information value (5.5) or (5.6) are relatively rare. For example, we have a change in sign of (5.5) for S_4 in the case of France and also in the case of Italy; but in 73 cases out of 8 × 10 = 80 there is indeed stability as to the sign. [Note that the sign of (5.5) for s = 1 is always opposite to that of s = 2, and similarly for (5.6) with s replaced by r.]

(2) The positive values of (5.5) for the exports of the E.E.C. countries (i = 1, ..., 5) to the E.E.C. (s = 1) apply to all commodity sets in all years with only 7 exceptions out of 150. The major exception is Italy's S_5. That country's exports of Animal and vegetable oils and fats to the E.E.C. is consistently below the independence

TABLE 10.12. MUTUAL INFORMATION VALUES (5.5) AND (5.6): E.E.C. VERSUS THE REST FOR INDIVIDUAL COUNTRIES BY COMMODITY SETS

Country	To E.E.C.			To Rest			From E.E.C.			From Rest		
	1961	1962	1963	1961	1962	1963	1961	1962	1963	1961	1962	1963
S_1: Food and live animals												
Belgium	.576	.572	.501 -	-1.703	-1.857	-1.542	.268	.262	.254 -	-.294	-.290	-.317
Netherlands	.344	.340	.319 -	-.690	-.725	-.719	-.664	-.941	-.992 -	.415	.524	.593 +
Germany	.405	.423	.382	-.883	-1.019	-.942	.611	.610	.616	-.934	-.944	-1.134 -
France	.427	.481	.470	-.960	-1.282	-1.360 -	.663	.713	.565	-1.080	-1.265	-.972
Italy	.290	.321	.285	-.544	-.668	-.612	.002	.311	.331 +	-.002	-.358	-.442 -
United Kingdom	.070	.169	.238 +	-.105	-.295	-.482 -	-.240	-.409	-.354	.187	.293	.290
Canada	-1.064	-1.429	-1.392	.796	.953	.986 +	-3.506	-3.771	-3.879 -	.867	.885	.966 +
U.S.A.	-.221	-.229	-.295 -	.264	.285	.371 +	-.749	-.765	-.739	.450	.460	.496 +
S_2: Beverages and tobacco												
Belgium	1.441	1.234	1.271	-5.469	-5.058	-5.674	.948	.724	.745	-.828	-.661	-.735
Netherlands	1.001	.839	.799 -	-1.222	-1.204	-1.041 +	.675	.575	.536 -	-.475	-.466	-.446 +
Germany	.172	.183	.318 +	-.108	-.147	-.271 -	.376	.375	.371 -	-.217	-.265	-.274 -
France	.222	.159	.215	-.144	-.126	-.170	-.170	-.049	-.006 +	.073	.027	.003 -
Italy	.692	.582	.493 -	-.623	-.636	-.479	.351	.210	.514	-.200	-.134	-.421
United Kingdom	-.943	-1.000	-.737	.349	.442	.353	-.295	-.154	-.251	.120	.080	.130
Canada	-4.591	-3.159	-3.527	.629	.711	.708	.160	.206	.119	-.082	-.131	-.075
U.S.A.	.364	.491	.270	-.258	-.496	-.222	-.581	-.673	-.650	.209	.275	.280 +

					S_3: Crude materials (inedible) except fuel											
Belgium	.665	.702	.704	+	-1.147	-1.206	-1.175		1.266	1.227	1.187	-	-1.302	-1.421	-1.411	
Netherlands	.669	.688	.726	+	-1.159	-1.160	-1.251	-	.500	.458	.533		- .278	- .276	- .356	
Germany	.780	.815	.819	+	-1.603	-1.674	-1.634		.786	.887	.799		- .527	- .729	- .648	
France	.849	.897	.881		-1.999	-2.182	-1.977		.390	.592	.659	+	- .203	- .391	- .480	-
Italy	.588	.577	.567	-	- .918	- .848	- .803	+	.651	.556	.617		- .398	- .358	- .436	
United Kingdom	.386	.492	.484		- .487	- .660	- .630		- .521	- .302	- .370		.174	.122	.151	
Canada	-2.243	-2.317	-2.526	-	.798	.784	.793		-3.890	-3.985	-4.045	-	.480	.524	.545	+
U.S.A.	.309	.245	.242	-	- .364	- .265	- .255	+	-1.870	-2.029	-1.873		.386	.436	.439	+
					S_4: Mineral fuels, lubricants and related materials											
Belgium	.497	.515	.367		-3.034	-2.803	-1.534	+	.546	.596	.523		-1.752	-1.627	-1.149	+
Netherlands	.177	.181	.145		- .472	- .448	- .393	+	.239	.149	.094	-	- .464	- .224	- .127	+
Germany	.526	.548	.485		-4.044	-3.692	-3.338	+	.045	.130	.267	+	- .072	- .191	- .426	-
France	- .179	.030	.003		.320	- .062	- .007		.595	.664	.570		-2.191	-2.139	-1.354	+
Italy	.102	- .003	.079		- .244	.007	- .196		- .238	- .396	- .731	-	.301	.398	.585	+
United Kingdom	.505	.545	.488		-3.230	-3.570	-3.455		.443	.616	.648	+	-1.151	-1.757	-1.799	-
Canada	-9.670	-10.202	-7.752		1.647	1.583	1.699		-8.660	-9.415	-7.459		1.337	1.214	1.170	-
U.S.A.	- .496	- .308	- .195	+	.697	.469	.362	-	-3.906	-4.777	-4.028		1.279	1.185	1.125	-
					S_5: Animal and vegetable oils and fats											
Belgium	.487	.476	.510		-3.883	-2.999	-3.465		1.143	.951	1.064		-1.356	-1.655	-1.704	-
Netherlands	.351	.293	.284	-	-1.469	-1.013	- .907	+	- .940	- .669	- .849		.312	.347	.359	+
Germany	- .237	- .096	.194	+	.435	.194	- .534	-	.151	.272	.278	+	- .083	- .238	- .209	
France	.409	.446	.425		-2.063	-2.402	-1.917		1.395	1.093	1.173		-2.486	-2.554	-2.308	
Italy	-1.334	-1.565	-1.230		1.263	1.311	1.157		- .500	- .598	.032		.199	.320	- .021	
United Kingdom	-1.085	-1.105	- .517		1.155	1.135	.719	-	- .876	- .882	-1.297	-	.298	.417	.462	+
Canada	-6.363	-2.468	-3.399		1.719	1.502	1.556		-3.626	-2.744	-3.436		.549	.698	.657	
U.S.A.	.141	.114	.044	-	- .393	- .293	- .098	+	1.269	.894	.862	-	-1.789	-1.427	-1.059	+

TABLE 10.12 (concluded)

Country	To E.E.C.			To Rest			From E.E.C.			From Rest		
	1961	1962	1963	1961	1962	1963	1961	1962	1963	1961	1962	1963
S_6: Chemicals												
Belgium	.285	.292	.367 +	-.574	-.617	-.949 -	.332	.336	.292	-.425	-.463	-.445
Netherlands	.341	.353	.314	-.741	-.811	-.750	.092	.034	.006 -	-.097	-.037	-.007 +
Germany	.256	.241	.206 -	-.499	-.477	-.426 +	.176	.195	.168	-.198	-.237	-.228
France	.123	.137	.181 +	-.207	-.243	-.364 -	.401	.452	.395	-.548	-.701	-.673
Italy	.197	.229	.174	-.359	-.449	-.345	.240	.217	.313	-.283	-.269	-.488
United Kingdom	.258	.232	.280	-.504	-.456	-.635	.215	.216	.165	-.248	-.268	-.223
Canada	-3.399	-3.317	-3.211 +	1.240	1.261	1.313 +	-2.894	-2.982	-3.107 -	.890	.936	1.033 +
U.S.A.	-.217	-.241	-.329 -	.276	.308	.421 +	-.028	-.098	-.178 -	.027	.096	.186 +
S_7: Manufactured goods classified chiefly by material												
Belgium	.398	.387	.364 -	-.710	-.720	-.749 -	.488	.425	.434	-1.195	-1.018	-1.145
Netherlands	.584	.583	.535 -	-1.362	-1.458	-1.453	.552	.540	.495 -	-1.520	-1.602	-1.470
Germany	.476	.461	.434 -	-.940	-.945	-.987 -	.417	.426	.400	-.911	-1.024	-.995
France	.523	.516	.470 -	-1.106	-1.147	-1.133	.505	.479	.461 -	-1.270	-1.259	-1.279
Italy	.171	.163	.198	-.241	-.238	-.337	.232	.269	.246	-.404	-.520	-.495
United Kingdom	.020	.044	-.012	-.025	-.059	.017	-.464	-.461	-.432 +	.470	.490	.490 +
Canada	-3.153	-3.662	-3.857 -	1.061	1.121	1.211 +	-2.261	-2.389	-2.378	1.075	1.135	1.177 +
U.S.A.	-.654	-.732	-.611	.532	.593	.573	-.663	-.673	-.763 -	.600	.633	.715 +

							S_8: Machinery and transport equipment									
Belgium	.655	.640	.563	−	−1.686	−2.085	−1.873		.571	.531	.477	−	−.960	−.953	−.918	+
Netherlands	.386	.212	.355		−.661	−.352	−.797		.530	.515	.509	−	−.848	−.904	−1.021	−
Germany	.285	.269	.258	−	−.439	−.474	−.511	−	.157	.229	.226		−.176	−.298	−.329	−
France	.459	.476	.451		−.862	−1.106	−1.181	−	.302	.263	.229	−	−.383	−.353	−.334	+
Italy	.340	.417	.374		−.555	−.885	−.863		.223	.274	.313	+	−.264	−.372	−.495	−
United Kingdom	.004	.030	.089	+	−.005	−.043	−.146	−	.375	.281	.196	−	−.508	−.383	−.277	+
Canada	−2.469	−2.426	−2.597		.988	1.080	1.183	+	−3.310	−3.298	−3.548		.927	.982	1.069	+
U.S.A.	−.671	−.723	−.841	−	.532	.623	.741	+	.113	−.024	−.055	−	−.122	.025	.063	+
							S_9: Miscellaneous manufactured articles									
Belgium	.652	.581	.504	−	−1.196	−1.116	−1.179		.415	.363	.337	−	−1.405	−1.222	−1.184	+
Netherlands	.715	.562	.509	−	−1.445	−1.050	−1.201		.396	.356	.349	−	−1.283	−1.182	−1.259	
Germany	.198	.235	.257	+	−.226	−.308	−.435	−	.241	.216	.141	−	−.598	−.564	−.355	+
France	.157	.110	.087	−	−.174	−.130	−.125	+	.219	.220	.193		−.528	−.576	−.525	
Italy	.111	.055	.007	−	−.118	−.062	−.009	+	.080	.056	.074		−.163	−.118	−.170	
United Kingdom	−.195	−.209	−.281	−	.170	.197	.307	+	.190	.140	.001	−	−.439	−.328	−.002	+
Canada	−2.083	−1.516	−.742	+	.810	.771	.621	−	−2.505	−2.412	−2.334	+	1.343	1.392	1.436	+
U.S.A.	−1.109	−.989	−.901	+	.612	.622	.697	+	.131	.080	.039	−	−.282	−.175	−.087	+
							S_{10}: Commodities and transactions not classified according to kind									
Belgium	.463	.492	.598	+	−.738	−.789	−1.078	−	.874	.773	.711	−	−1.819	−1.451	−1.030	+
Netherlands	.423	.495	.716	+	−.645	−.796	−1.550	−	.622	.659	.573		−.900	−1.054	−.714	
Germany	.529	.476	.548		−.914	−.748	−.927		.195	.316	.573	+	−.192	−.357	−.715	−
France	.402	.533	.496		−.599	−.896	−.787		.643	.676	.784	+	−.952	−1.104	−1.246	−
Italy	.356	.705	.528		−.507	−1.523	−.870		−.009	.274	.177		.008	−.298	−.160	
United Kingdom	−.018	−.048	−.036		.019	.048	.036		.913	.938	.969	+	−2.057	−2.475	−2.102	
Canada	−2.184	−2.842	−3.382	−	.868	.919	.946	+	−3.767	−3.944	−4.110	−	.846	.877	.811	
U.S.A.	−1.509	−1.775	−1.873	−	.754	.793	.803	+	−.243	−.359	−.735	−	.181	.260	.400	+

level, and hence the exports to the Rest above that level.

(3) There are more exceptions to the positive sign of (5.6) for the imports by the E.E.C. countries ($j = 1, \ldots, 5$) from the E.E.C. ($r = 1$). The number of negative signs is 15 out of 150. The Netherlands is a major exception with respect to S_1 and S_2: Its imports from the E.E.C. of Food and live animals and of Animal and vegetable oils and fats are almost 1 bit below the independence level.

(4) The values of (5.5) for the exports of the United States to the E.E.C. are negative except for 9 positive signs out of 30. Beverages and tobacco and Crude materials (inedible, fuel excepted) are the main exceptions; for these the exports to the E.E.C. are above the independence level. The values of (5.6) for the U.S. imports from the E.E.C. are also mostly negative. The number of exceptions is 7 and the major exception is Animal and vegetable oils and fats.

(5) All Canadian values of (5.5) for its exports to the E.E.C. are negative, and many of them are substantial in absolute value. Those of (5.6) for imports from the E.E.C. are negative too except for Beverages and tobacco.

(6) As could be expected, the picture of the United Kingdom is more diffuse than that of any other country: 17 positive and 13 negative values for (5.5) in the case of exports to the E.E.C., 15 positive and 15 negative values for (5.6) in that of imports from the E.E.C.

(7) As in Table 10.11 each of the three-element time series is followed by + or by - except when the development is not algebraically monotonic. The percentage of nonmonotonic cases (no + or -) is almost 50. In Table 10.11 it is only 20. This difference reflects the increase in erratic behavior when the object of analysis is more microeconomic in nature. In this case: when we shift from total trade between countries and country sets to the trade in particular commodity sets. A still more detailed analysis (commodities rather than commodity sets, or country-to-country trade rather than country-to-country set trade) would undoubtedly be interesting from the standpoint of the detail involved, but the decrease in regularity would almost surely continue.

APPENDIX TO CHAPTER 10

10.A. The RAS Method for Adjusting Bivariate Frequency Tables

The RAS method,[1] when described in the terminology of Section 10.1, amounts to the following. One multiplies each x_{ij} by a number r_i which is specific for the exporting region and by a number s_j which is specific for the importing region. These numbers are to be chosen such that the marginal constraints are satisfied. Hence:

$$y^*_{ij} = r_i x_{ij} s_j \qquad i, j = 1, \ldots, n \tag{A.1}$$

$$\sum_{j=1}^{n} y^*_{ij} = r_i \sum_{j=1}^{n} x_{ij} s_j = y_{i.} \qquad i = 1, \ldots, n \tag{A.2}$$

$$\sum_{i=1}^{n} y^*_{ij} = s_j \sum_{i=1}^{n} r_i x_{ij} = y_{.j} \qquad j = 1, \ldots, n \tag{A.3}$$

where y^*_{ij} is the RAS forecast of y_{ij}. [The word RAS is due to the original use of the method in input-output analysis; the input coefficients, usually denoted by a_{ij}, are the object of adjustment and the adjusted value is of the form $r_i a_{ij} s_j$.]

The economic interpretation of the procedure is as follows. When the share y_{ij} of the exports from i to j exceeds the earlier value x_{ij}, this may be due either to the fact that the exports of i show an overall increase, or to the fact that the imports of j show a general increase, or to the fact that there is a special increase of the exports from i to j which is neither related to i's total exports nor to j's total imports. It is assumed that there is no such special effect, in which case it is reasonable to say that r_i measures the effect of the change in the share of i in total exports ($r_i > 1$ if the change is positive, $r_i < 1$ if it is negative), and that s_j measures in an analogous way the effect of the change in j's share of total imports.

If we substitute in (A.1) the ratios $y_{i.}/x_{i.}$ and $y_{.j}/x_{.j}$ for r_i and s_j, respectively, we find that y^*_{ij} is equal to y'_{ij} as defined in (1.4). They are not really equal, of course, because y^*_{ij} does and y'_{ij} does not satisfy the marginal constraints. But if we adjust the y'_{ij} such that these constraints are satisfied, it stands to reason that the result ($\hat{y}_{ij}$) will not differ very much from y^*_{ij}. Hence the RAS forecast and the two-stage information forecast will be approximately equivalent as long as the y'_{ij} of (1.4) do not violate the marginal

[1] See R. Stone and A. Brown, A Computable Model of Economic Growth. Volume 1 of the series "A Programme for Growth" (London: Chapman and Hall, Ltd., 1964).

constraints too seriously. In fact, there are no appreciable differences between the results of the two methods when applied to the data of Section 10.2; reference is made to the article by Uribe, De Leeuw and Theil quoted at the beginning of this chapter.

10.B. Some Details on the Import-Export Data of Eight Countries

The data analyzed in Sections 10.3 through 10.5 are taken from the O.E.C.D. Statistical Bulletins on Foreign Trade, Series C: Trade by Commodities. All flows are measured in units of $1000. The 183 commodities are those three-digit groups for which the transactions are sufficiently large; all groups for which transactions exceed $10,000 are specified separately, and the remainder is summarized under one of the three-digit groups of the last one-digit group (S_{10}).

For 1963 there are no complete data on Dutch exports with respect to composition. For the purpose of the computations described in the text these remainder items have been allocated proportionally over all commodities. The size of these items, expressed as a percentage of Dutch exports to the seven other countries in 1961, is as follows: Belgium 7.4; Germany 4.8; France 7.2; Italy 12.4; United Kingdom 9.4; Canada 21.3; U.S.A. 16.3.