THE VALIDITY

OF

INTERNATIONAL

GOLD MOVEMENT

STATISTICS

OSKAR MORGENSTERN

This is the second number in the series called SPECIAL PAPERS IN INTERNATIONAL ECONOMICS, *to be published from time to time by the International Finance Section of the Department of Economics and Sociology in Princeton University.*

This series was inaugurated recently to include important work of less than book length whose publication would be in the interest of many of those concerned with international affairs but which, for one reason or another, does not fit gracefully into the journals or into either of two other series sponsored by the Section: ESSAYS IN INTERNATIONAL FINANCE *and* PRINCETON STUDIES IN INTERNATIONAL FINANCE.

The writers of these papers are free to develop their topics as they will and sponsorship by the Section carries with it neither approval nor disapproval of the ideas expressed.

As with the ESSAYS, *it is planned to distribute with our compliments single copies of these pamphlets to persons showing an interest in them.*

<div align="right">

GARDNER PATTERSON, *Director*
International Finance Section

</div>

THE VALIDITY

OF

INTERNATIONAL

GOLD MOVEMENT

STATISTICS

OSKAR MORGENSTERN

CONTENTS

CONTENTS

TABLES AND CHARTS

THE AUTHOR

Oskar Morgenstern is Class of 1913 Professor of Political Economy at Princeton University. Before coming to Princeton in 1938 he was Professor at the University of Vienna and Director of the Austrian Institute of Business Cycle Research. He was a Member of the Committee of Statistical Experts, and of the Delegation on Economic Depressions, both of the League of Nations. He is now directing, at Princeton University, the Economics Research Project on the Mathematical Structure of American Type Economies, supported by the Office of Naval Research.

Among his publications are: *Theory of Games and Economic Behavior* (with J. von Neumann), *On the Accuracy of Economic Observations, Economic Activity Analysis,* and *International Financial Transactions and Business Cycles* (in Press).

I. Introduction

It is the purpose of this paper to examine the statistics of international gold movements for their reliability and usefulness. An investigation of this kind is justified for the light it may shed on the problem of analysis of economic data in general, and, specifically, because gold movements are of great importance in the study of international trade, especially the behavior of exchange rates and capital movements. Gold movements are of particular significance for our understanding of the period of the gold standard, but they are also important under other monetary systems. Their accuracy, it should be noted, is probably one of the least questionable of all the components of balance of payments statistics, certainly not to be compared with such items as security movements, the pure estimates of tourists' expenditures, and immigrants' remittances. It is of singular importance to form a good idea of the worth of one of the most trustworthy components. In the absence of fortuitously offsetting errors, the aggregate, formed of parts of widely diverse quality, cannot be better than the best part—in fact, it will in almost all cases be considerably below that standard.[1]

A study of the statistics of international gold movements is a part of the larger question of the accuracy of economic observations in general and of those of trade in particular.[2] But in the

[1] The study described in the following pages grew out of the work involved in a larger investigation that has been undertaken for the National Bureau of Economic Research, to be published in the near future by Princeton University Press under the title *International Financial Transactions and Business Cycles*. In the course of this investigation, gold movements, gold points, their relations to exchange rates, etc. had to be examined. It was hoped that significant correlations could be computed, on the basis of monthly data, for exchange rates and gold movements for the important countries of the world during a period from approximately 1870 to 1939. However, studies by other writers of these relationships had already shown that it was very difficult, if possible, to obtain intelligible and significant relationships. The question therefore arose whether there might be troubles with the basic statistics themselves. Few difficulties are to be expected or found in the statistics of exchange rates. Statistics of gold points are non-existent; a collection of all accessible information about gold points for New York, London, Paris, and Berlin from 1870-1938 is made in Chapter V of my above-mentioned forthcoming book. The plentiful statistics of gold movements lent themselves to an investigation of their accuracy.

[2] See my book, *On the Accuracy of Economic Observations*, Princeton: Princeton University Press, 1950; second revised edition, 1956.

case of gold, one is in a particularly favorable position because one is dealing with a clearly defined commodity of great significance. There are no technological changes in the product considered, and there is no serious problem of classification; it is easily ascertained what constitutes gold, though not necessarily monetary gold. Furthermore, the commodity itself has been in international trade without interruption for many centuries. Moreover, the statistics of gold movements have been used by hosts of writers in their analyses of the conditions in various countries, in the studies of balance of payments, and for the verification and development of the theory of money and international trade. Many of the observations made by these authors ultimately go back either explicitly or implicitly to information given by the statistics of gold movements across national borders. It is not necessary here to elaborate on the significance of such writings;[3] it suffices to state that most observations regarding gold in the field of the theory of international balances ultimately rely on information that can be culled only from that given by international gold movement statistics. Consequently, a test of the value and the significance of these statistics is of considerable interest.

The "test" described in the next pages is entirely on a common-sense basis. There is—apart from some measures in Table 3, which involve nothing deeper than the notion of a standard deviation in a sample test—no device or argument that could not have been used a hundred years ago, or even in the times of Cantillon and Adam Smith. The fact that no high-powered statistical procedures are used is explained by the peculiar circumstance that no techniques, specifically appropriate for the subsequently shown situation, exist, the tremendous development of statistics during the last decades notwithstanding. Yet for a first attempt this is not an insurmountable difficulty; it is not necessary to use advanced scientific techniques unless the latter carry us very much further in a given situation than common-sense exploration. Should spe-

[3] See Carl Iversen, *Aspects of the Theory of International Capital Movements*, Copenhagen and London, 1935, for a rather comprehensive survey of the literature. See also the standard works by Angell, Haberler, Ohlin, and Viner. Iversen does not raise the problems of the validity of the statistics, although he quotes H. D. White's book which notes some of the difficulties, nor does the question come up in the most recent book specifically devoted to gold movements: W. E. Beach, *British International Gold Movements and Banking Policy, 1881-1913*, Cambridge, Mass., 1935.

cial statistical methods upset our common-sense analysis in a new and unfamiliar field, we would, in fact, be justified in rejecting them until there is much corroboration of the results obtained with the special devices.

In order to determine the value of statistics of gold movements, it is necessary to establish a *standard of accuracy*. This standard, by common scientific practice, can only be derived from the uses to which the data are to be put. The uses are, of course, governed by the theories for which the data are needed. If no theories existed, it would be impossible to say what accuracy of measurement of these quantities means. It will be seen in the following, however, that even for ordinary description, which is in itself not intended to be an application of high-powered economic theory, the statistics are so grossly bad as, for most purposes, to be rejected.

Data on international gold movements are used not only for purposes of description and analysis, but also for policy guidance. The rates of interest of central banks, for example, frequently have been changed because of reported gold flows between countries. Such changes were justified by data which—as we shall see—are of highly questionable character even for these purposes.

There is no doubt that all serious economic writers try to be as careful about their sources as possible, that they are constantly looking for the best sources, and that they take great pains to collate the transcription of their data. Yet all this is preliminary. The work is only properly done when a numerical estimate of the error of the observations is provided to the reader. Lacking this numerical measure, a full qualitative description of the intuitively knowable errors is indicated, together with restricting the use of the data beyond the first few significant digits.

Whatever the final verdict about the gold movement statistics may turn out to be, they do illustrate a deep and peculiar difficulty common to all economic investigations: It is necessary to combine, in interpretations, quantitatively given data such as these with qualitative, historical descriptions of phenomena from which the data arise and of which they are only a part. Unquestionably, these qualitative descriptions also contain strong elements of doubt, inaccuracy, and error which are even less amenable to precise

3

formulation. The problem is to determine in which of these components the errors are larger and how they may interact upon each other. In this study we deal only with the quantitative component, the larger question going far beyond the compass of our investigation. We finally state that data found to be of poor quality for the past need not remain so for the future.

II. Uses of Gold Movement Statistics

The statistics of international gold movements are used primarily for two general purposes: (1) broad observations of long-run behavior of, say, gold and prices, using years as the minimum unit, and (2) the study of the *fine structure* of international economic interaction, as described, for example, by the variants of the theory of the gold standard and various theories of adjustment of balances of payments. For the latter, monthly data are indispensable.[4] Obviously, the criteria of the usefulness of the data must be taken from the *maximal requirements* to be made for them. At the same time, it must be emphasized that the maximal requirements set in this study are not different from those used in the past and that none exceed those which have been made, tacitly or openly, by persons using such statistics, or should have been made had the subject even been touched on by these writers.[5]

Gold statistics fall into two main parts: production and movements. Both are domestic and international. We are here concerned only with gold movements across national borders. If monetary gold movements could be separated from industrial—in all respects—this would be highly desirable since the indirect effects of gold as an industrial commodity upon the price system and upon the international accounts are much less important and less direct than its monetary effects.

The separation of monetary and non-monetary gold is neither simple nor conclusive. Gold can move from one category into the other even within one country and domestic gold production can affect the stocks of both. During the classical gold standard period it was impossible to know, in the vast majority of cases, whether gold leaving and arriving came from one or the other of these sources and whether it was going—or in which proportions—to industrial or monetary use. If monetary and non-monetary gold

---

[4] If practicable, one might choose weekly or daily data as even better fitting the ideas of the underlying mechanisms.

[5] There are important and widely quoted empirical investigations in which the question of the accuracy of the gold movement data is not once mentioned! The data are often subjected to elaborate calculations—and similarly far-reaching conclusions are drawn.

movements are separated by the reporting countries, the criteria for separating them should be made known for it is impossible to disentangle the reports once they are published. Instead of merely reporting as accurately as possible, the statistical agencies can also "interpret" what they report, thereby possibly confusing matters still further. For example, it can be argued that receipt by the central bank of domestically mined new gold is equivalent to a commodity export of the country because it increases the country's holdings of gold or foreign exchange. Conversely, gold consumed at home in industry or the arts has the same effects as a commodity import. If this practice is used, we have "international" transactions which involve only one country. Further difficulties arise when the amount of gold content in industrial articles is considered, such that, for example, commodities with 80 per cent or more gold content are classified as equivalent to bullion.[6]

From the statistics of monetary gold movements plus those of some[7] gold stocks (in banks and treasuries) all other gold statistics are derived. On them hinge important judgments, including the nature and extent of the "gaps" in the balance of payments. Information about gold in possession of central and commercial banks and of treasuries is generally made available not only in the usual trade statistics but also in special reports. Nonetheless, it is seldom trustworthy. At any rate, none of the available statistics shows the variations in banks' and treasuries' possessions as originating from contact with specific single countries; they are thus

[6] Some of these questions are discussed by W. R. Gardner, "Merchandise Trade in the Balance of Payments" in *International Trade Statistics*, edited by R. G. D. Allen and J. E. Ely, New York, 1953 (for some comments on this, see my review in *Econometrica*, Vol. 23, 1955, pp. 105ff.). This important work shows many of the tremendous difficulties of obtaining correct foreign trade statistics, but it offers no method for the numerical estimation of the errors. Gardner's discussions (the International Monetary Fund's procedures which he describes) are incomplete in one important respect: If gold leaves *old* hoards, it is equal in its *monetary function* to newly mined gold and should be treated accordingly; if the hoards are of shorter duration than is necessary for the economy to absorb all effects of the creation of the hoard, there is no such equivalence. How are the Fund statisticians going to take care of this dilemma, which may be quantitatively very important, and, incidentally, in its wider aspects, ought to be resolved by monetary and business cycle theory?

[7] Not all. Gold hoards are of unknown but very great magnitude in certain countries, for example, India. Gold movements comprise also some, but not all, changes of hoards.

only useful for more limited investigations than those now under discussion.

Every important country shows in its foreign trade statistics transactions of gold and silver, sometimes in great detail. These official records are the ultimate sources of information on movements; all other data are compilations from them. Reports in trade journals are either taken from these or they are based on hearsay; trustworthy continuous series cannot be constructed from them. As always, when foreign trade statistics are involved, a simple though laborious *test* of the data is possible and indicated. It is the one developed (for gold) by C. F. Ferraris.[8] His execution of the program was modest but in principle correct. Ferraris measured the movement of gold from France to Great Britain for the periods 1876-1880 and 1881-1884 according to the statistical sources of both countries. The result was that English import statistics showed in each of these periods a volume more than double that given in the French export statistics.[9] As in other instances when the untrustworthiness of data was shown, there is no evidence that Ferraris' contemporaries, in using gold statistics, paid much attention to his results which, in fact, should have caused not only caution in the use of existing information but also have led to reforms in making new statistics. The neglect of these observations is the more remarkable since Ferraris was neither the first nor the last to point out difficulties encountered with statistics of gold movements. More famous writers, such as A. Soetbeer, de Foville, R. Giffen, and others have, from time to time, commented in the same vein.[1] It is indeed peculiar that there has been

[8] *La Statistica del Movimento dei Metalli Preziosi fra l'Italia e l'Estero*, Rome, 1885. This small book has been almost wholly neglected.

[9] See also Ferraris' contributions (in French) to *Bulletin de l'Institut International de Statistique*, Vols. 11 and 12 (1899-1900), especially Vol. 12, pp. 338-340.

[1] J. Viner mentions, in his *Canada's Balance of International Indebtedness 1900-1913*, Cambridge, Mass., 1924, p. 33, that even according to Canadian sources their statistics of movements of gold *coins* are imperfect. Clearly, these are more difficult to make than those for movements of bullion and industrial gold. J. H. Williams: *Argentine International Trade under Inconvertible Paper Money 1880-1900*, Cambridge, Mass., 1920, does not concern himself specifically with this property of his data and uses Argentinian gold figures to 8 significant digits—a dangerous undertaking, as the subsequent pages will demonstrate, especially when one considers that Argentinian statistics of that period were in general hardly superior to those of Britain, Germany, and many other European nations. W. E. Beach: *British International Gold Movements and Banking Policy, 1881-1913*, Cambridge, Mass., 1935, is a generally painstaking work. The question of the validity

no reform of the statistics produced by the official agencies, and that later scholars have insisted on using the data in countless instances as if they were reliable.[2]

The test in the following section uses Ferraris' principle and consists in taking A's exports to B according to A's export statistics and B's import statistics and B's exports to A according to B's export statistics and A's import statistics. Each of these two sets should agree within certain bounds that are set by such items as freight, insurance, and duties. To these some other factors may have to be added (see below, page 32ff.), mostly in the nature of lags; a reasonable allowance for clerical and printing errors must also be made.

In anticipation, it may be said that the test establishes that the statistics of pairwise gold movements, and *a fortiori* those derived from them, are worthless from the point of view of describing the *finer* interdependencies among national economies. By implication only very few uses of such highly inaccurate material are justifiable in economics. This result is a great disappointment because it means that much economic reasoning in an important field has been based on guesswork as to the underlying empirical data.

---

of gold statistics is not raised, except for comparing Customs House and Bank of England statistics, and very ambitious analyses are made even of such things as the drain of gold into domestic circulation—all this to many significant digits. If proper warning flags had been hoisted, the reader would have benefited.

[2] An interesting case is H. D. White: *The French International Accounts 1880-1913*, Cambridge, Mass., 1933, who, on pp. 43-48, discusses the low value he sees in the statistics. He enumerates some of the literature, also apparently quoting one of Ferraris' minor writings, though misspelling his name consistently. Yet, in later parts of his book, the net movements of specie for France are minutely compared with other time series as if there were nothing to worry about.

III. The Test of Validity

The procedure used for testing the validity of gold statistics is described in some detail so as to make easily possible a test for other countries and periods. At the outset it must be recognized that our conclusions are, to some extent, a matter of individual judgment, and our opinions may of course be rejected by others. It would be hard to choose between such conflicts in opinion because—as will be seen—statistical problems are involved for which there exists at present no statistical theory which one would wish to fall back upon. The entire case is of value, however, for all students of economics, for it shows in an important field what precautions must be taken before establishing theories or "verifying" them.

The problem may best be broken down into four parts: (1) the formulation of the underlying hypothesis, (2) the description of the sample and the results, (3) the explanation of the failure of the statistics to confirm the hypothesis and a discussion of certain difficulties even if it had been confirmed, and (4) the consequences of (2) and (3).

(1) *Formulation of the Hypothesis*:

The hypothesis is rather easily formulated. The observation of gold movements from country A to country B are made in A's export and in B's import statistics. The two should coincide, allowing for such factors as costs of transportation and time lags. In neighboring countries both virtually disappear; in countries more distant from each other, say Europe and Australia, a time lag may have to be considered. Costs of transportation, insurance, handling charges, and so on, which all enter significantly in the determination of gold points, play no appreciable role here. Gold is so valuable that these costs all but disappear from the statistics, as is evident from the proportion they make with the value of gold shipments.[3] In view of the rounding off of numbers in foreign trade statistics, the factor of distance (costs of transportation)—which plays a great role in determining the value of less valuable

[3] This does not contradict the fact that the *profitability* of shipments depends on very small changes of transportation costs, interest rates, etc.

commodities, and, consequently, is a source of discrepancies in uncorrected statistics[4]—can often be neglected in the case of gold statistics. We shall nevertheless not expect a 100 per cent agreement in the two statistics describing the same foreign trade operation, but will accept "reasonable" deviations. However, differences also due to other causes should be random. If they are not, there is probably a special factor or bias that should be accounted for separately.

Thus, the expectation of practical identity of the two sets of data for neighboring countries after allowing for a possible, but not necessary, time lag for others, constitutes the hypothesis of the following test.

(2) *Description of the Sample Test and Its Results*:

(a) *The data* used are the official statistics of imports and exports of gold, gold coin, bullion, leaf, etc., behind which it is impossible to go. The sources and further particulars are indicated in the appendix containing footnotes and sources for the data in Tables 1 and 2. Sterling, instead of dollars, was chosen as the unit of measurement because most references to gold movements are in sterling, although the test, determining the ratio between the two observations, is independent of units and could have been made in any currency. In addition to the four principal countries, United Kingdom, United States, Germany, and France, Canada has been included for two reasons: first, because it is desirable to have more than one, and indeed adjoining, countries in North America; and second, because Canada is of great significance in the trade, being among other things, a country through which gold was often routed when it moved between the United States and the United Kingdom. Transit is a major complication, as we shall presently see.

(b) *The period* covered by the sample is the years 1900, 1907, 1928, and 1935, two before and two after World War I in equal distances from each other. However, Charts 5 and 6 show monthly time series over longer consecutive periods because it seemed important to give some longer time series, demonstrating that the choice of one source rather than of another (both pre-

---

[4] It leads to the popular "paradox" that the recorded total of world exports is less than the total of world imports.

sumably equivalent) leads to widely different results concerning cycle turning points and the duration of cycles. This could not be accomplished by restricting the study to the sample years alone.

It is unnecessary to go back before 1900 since one would expect the reliability of the data to be even lower then and because we have the general indications noted earlier of Ferraris and de Foville for France and England, which confirm this expectation. 1900 was a year when the business cycles in the three European countries were at a peak, while the American economy was contracting, reaching its trough in December; 1907 was a year of great crisis with violent movements; 1928 was an upper turning point for business in most European countries (including the three concerned); and 1935 was one of moderate activity under conditions of widespread government, or central bank, supervision of formerly free gold movements. It can, then, be said that these four years show both a sufficient variety and similarity of cyclical pattern. Gold movements of many sorts and dimensions could be expected in these widely different periods to the exclusion of any particular bias. The period after World War II was not considered because it is still too short for significant additions to time series and because the anomalies of the transactions in these years prevented gold movements from playing a principal role in basic studies of the theory of international trade and money.

For each year *monthly data* were assembled from both sources in a total of eight directions. Canada was introduced only in her relations with the United Kingdom and the United States. This gives six pairs of countries plus two pairs involving Canada. Since each country's export and import statistics were examined and compared with the import and export statistics of each other country, we deal, for the four chosen years, with a possible 1536 months.[5] However, since only the *differences* between the reports enter into our subsequent computations, the number is cut to one-half, or 768 months. This is a large sample amounting to 14.8 per cent of the total possible number of months for 1900-1914 and 1925-1938. It is big enough to use with confidence if the answer looked for is sufficiently clear; it is even satisfactory if we are searching for a view about the period before 1900, when the

---

[5] This number is hypothetical because there were in fact somewhat fewer entries inasmuch as data were not available for each country for every single month.

11

TABLE 1

Comparison of Annual Gold Movements and Ratio of Pairs of Statistics for Selected Countries for 1900, 1907, 1928 and 1935[1]

Unit: Thousands of Pound Sterling

		U. K.-Germany	Germany-U. K.	Ratio Germany to U. K.[3]	U. K.-France	France-U. K.	Ratio France to U. K.	U. K.-U. S.	U. S.-U. K.	Ratio U. S. to U. K.	Germany-U. S.	U. S.-Germany	Ratio U. S. to Germany
1900	Exports[2]	4,903	2,065	1.12	2,360	671	1.76	1,082	3,896	.68	—	1,052	—
	Imports	2,544	5,490	.81	2,154	4,165	.31	5,872	738	.66	804	776	1.31
1907	Exports	3,162	7,176	1.25	7,208	3,224	.92	18,354	898	.99	210	558	2.25
	Imports	7,760	3,947	.93	5,204	6,590	.62	652	18,104	1.38	—	473	—
1928	Exports	21,047	116	—	19,969	—	.01	6,602	6,684	1.17	19	5,913	—
	Imports	52	—	2.22[4]	187	285	—	6,684	7,708	1.00	—	—	—
1935	Exports	1,373	149	—	49,845	77,164	1.00	85,883	117	.75	10	36	.80
	Imports	454	—	.33	81,000	49,631	.95	168	64,485	.70	—	8	—

		Germany-France	France-Germany	Ratio France to Germany	U. S.-France	France-U. S.	Ratio France to U. S.	U. K.-Canada[5]	Canada-U. K.	Ratio Canada to U. K.	U. S.-Canada	Canada-U. S.	Ratio Canada to U. S.
1900	Exports	16	—	28.1	3,527	185	.47	—	—	—	645	204	.87
	Imports	130	450	—	368	1,667	.50	—	21	—	1,387	558	.15
1907	Exports	481	—	1.22	5,197	422	.94	—	1	—	2,795	3,955	.47
	Imports	641	589	—	1,035	4,879	.41	—	135	—	5,197	1,309	.76
1928	Exports	—	—	—	63,296	—	—	—	1,958	—	4,653	21,323	1.04
	Imports	—	—	—	31	—	—	1,958	1,014	1.00	21,038	4,854	1.01
1935	Exports	18	1,832	15.4	12	186,791	84.91	—	2,164	—	15	17,466	11.20
	Imports	—	277	—	192,005	1,019	.97	2,174	1	.99	19,573	168	.89

assumption is made that the statistical data then cannot, in general, have been more trustworthy than after 1900. In view of the size of the sample, it is not necessary to go into details of modern sampling theory, which only assume significance when the numbers are considerably smaller.

From the monthly data—which are not reproduced except in part in Charts 5 and 6, to be discussed later—the quarterly and yearly values were obtained by summation. The monthly values were necessary to check on lags and, in that connection, for the formation of cumulative sums shown in Charts 4, 7, and 8.

(c) *The monthly* statistics are summarized in Table 1 in annual form and in Table 2 in quarterly form.[6] To them correspond Charts 1, 2, and 3.

Each table, 1 and 2, gives exports and imports of gold in 1000 pound sterling from both sources, where available. There is also shown the *ratio* between the statistics from the two sources. If both sets of data were accurate, the ratios should be unity or very near to unity.[7] The results shown are extraordinary. Not only do the ratios not cluster closely around one, as they should, but quite absurd and fantastic values occur. The spread, as is to be expected, is greater for the quarterly than for the yearly data and would still be larger for monthly data were they shown.

An even more comprehensive picture is obtained from Charts 1, 2, and 3, essentially frequency distributions. These charts also suggest the possibility of forming and evaluating an additional hypothesis which might improve the meaning of the results. On the horizontal axis the ratios are shown and on the vertical axis the

[6] Monthly data are not shown because of their bulk and in order to avoid any impression of erratic behavior of the sample. See, however, the monthly data of net movements in Charts 5 and 6.

[7] How near unity is a problem for which a statistical theory would be required (see below).

Notes to Table 1:

[1] The figures are those of the first named country for each pair. In some instances the sum of the monthly figures does not add up to the annual total given in the trade statistics. When this occurs the final figure reported, that is, the annual total, is used here. A dash means negligible amounts. For sources and notes see Appendix.

[2] The figures should be compared on the diagonal for each pair.

[3] The ratio is the proportion of imports from A to B to exports from A to B; e.g.:

$$\frac{\text{B's imports from A in B's books.}}{\text{A's exports to B in A's books.}}$$

[4] Pairs of data with italicized ratios are excluded from the charts due to the smallness of the original figures or to the large size of the ratio.

[5] The pairs including Canada cover only July to December for 1900.

TABLE 2

Comparison of Quarterly Gold Movements and Ratios of Pairs of Statistics for Selected Countries for 1900, 1907, 1928 and 1935[1]

Unit: Thousands of Pound Sterling

Quarter	U. S. Exports to France[2]	France Imports from U. S.	Ratio[3] France Imports to U.S. Exports	U. S. Imports from France	France Exports to U. S.	Ratio France Exports to U.S. Imports	U. K. Exports to Germany	Germany Imports from U. K.	Ratio[3] Germany Imports to U.K. Exports	U. K. Imports from Germany	Germany Exports to U. K.	Ratio[3] Germany Exports to U.K. Imports
1900												
1st	—			116			5	1	.20	1,764	1,574	.89
2nd	2,575	No	No	5	No	No	851	59	.07	520	325	.62
3rd	952	data	data	—	data	data	1,133	1,495	1.32	185	140	.76
4th	—			247			2,914	3,935	1.35	75	26	.35
1907												
1st	—	6	—	68	16	.24	4	221	55.25	476	103	.22
2nd	4,429	376	.08	—	—	—	28	747	26.68	182	72	.40
3rd	728	4,439	*6.10*	79	59	.75	2,178	1,308	.60	82	71	.87
4th	40	58	1.45	888	347	.39	952	1,671	1.76	7,020	6,928	.99
1928												
1st	10,439			2			402			1	30	*30.00*
2nd	40,229	No	No	23	No	No	290	No	No	2	27	*13.50*
3rd	12,628	data	data	3	data	data	7,132	data	data	48	29	.60
4th	—			3			13,222			1	30	*30.00*
1935												
1st	3	141	*47.00*	12,418	13,213	1.06	8			32	26	.81
2nd	—	832	—	72,104	68,154	.95	210	No	No	116	23	.20
3rd	9	3	.33	8,324	4,847	.58	934	data	data	46	49	1.07
4th	—	43	—	99,159	100,577	1.01	214			260	51	.20

TABLE 2 (continued)

Quarter	U.K. Exports to U.S.	U.S. Imports from U.K.	Ratio[3] U.S. Imports to U.K. Exports	U.K. Imports from U.S.	U.S. Exports to U.K.	Ratio[3] U.S. Exports to U.K. Imports	U.K. Exports to France	France Imports from U.K.	Ratio France Imports to U.K. Exports	U.K. Imports from France	France Exports to U.K.	Ratio[3] France Exports to U.K. Imports
1900												
1st	2	16	*8.00*	2,578	751	.29	56			477		
2nd	6	10	*1.67*	3	—	—	484	No	No	1,189	No	No
3rd	—	—	—	3,281	3,141	.96	600	data	data	392	data	data
4th	1,074	712	.66	10	4	.40	1,220			96		
1907												
1st	155	172	1.11	29	—	—	281	345	1.23	1,377	1,321	.96
2nd	398	408	1.03	22	568	*25.82*	3,804	3,149	.83	326	210	.64
3rd	—	14	—	585	206	.35	826	925	1.12	80	99	1.24
4th	17,801	17,510	.98	16	124	7.75	2,297	2,171	.95	3,421	1,594	.47
1928												
1st	—	—	—	5	210	*42.00*	18,660	10	—	1		
2nd	—	—	—	6,677	6,474	.97	968	18	.02	—	No	No
3rd	500	500	1.00	—	—	—	146	99	.68	18	data	data
4th	6,102	7,208	1.18	2	—	—	195	158	.81	168		
1935												
1st	30,964	30,412	.98	72	95	1.32	10,298	12,826	1.25	13,240	8,679	.66
2nd	7,432	838	.11	61	—	—	31,757	30,330	.96	57,784	55,880	.97
3rd	18,713	7,579	.41	1	—	—	5,024	5,953	1.18	1,748	5,836	*3.34*
4th	28,774	25,656	.89	34	22	.65	2,766	679	.25	8,228	6,769	.82

[1] A dash means negligible amounts. For sources and notes see Appendix.
[2] In each pair the statistics are those of the first country named.
[3] Pairs of data with italicized ratios are excluded from the charts due to the smallness of the original quarterly figures or to the large size of the ratio.

TABLE 2 (continued)

Comparison of Quarterly Gold Movements and Ratios of Pairs of Statistics for Selected Countries
for 1900, 1907, 1928, and 1935[1]
Unit: Thousands of Pound Sterling

Quarter	U. K. Exports to Canada[2]	Canada Imports from U. K.	Ratio Canada Imports to U. K. Exports	U. K. Imports from Canada	Canada Exports to U. K.	Ratio Canada Exports to U. K. Imports	U. S. Exports to Canada	Canada Imports from U. S.	Ratio[3] Canada Imports to U. S. Exports	U. S. Imports from Canada	Canada Exports to U. S.	Ratio Canada Exports to U. S. Imports
1900												
1st		No					92	No	—	764	—	—
2nd	No	data	No	No	No	No	453	data	—	418	—	—
3rd	data	1	data	data	data	data	453	237	.52	542	154	.28
4th		20					192	321	1.67	845	50	.06
1907												
1st		34					326	202	.62	1,005	1,269	1.26
2nd	No	4	No	No	No	No	1,127	344	.31	731	860	1.18
3rd	data	42	data	data	data	data	889	293	.33	946	242	.26
4th		55			1		453	470	1.04	2,515	1,584	.63
1928												
1st		—		750	750	1.00	34	120	3.53	9,485	10,656	1.12
2nd	No	—	No	229	229	1.00	23	86	3.74	4,181	2,284	.55
3rd	data	500	data	—	—	—	30	157	5.23	2,206	3,692	1.67
4th		514		979	979	1.00	4,566	4,491	.98	5,166	4,691	.91
1935												
1st		—		618	698	1.13	6	28	4.67	4,503	4,834	1.07
2nd	No	1	No	786	700	.89	6	30	5.00	4,682	4,064	.87
3rd	data	—	data	365	365	1.00	3	59	19.67	5,396	5,247	.97
4th		—		405	401	.99	—	51	—	4,992	3,321	.67

TABLE 2 (continued)

Quarter	German Exports to France	French Imports from Germany	Ratio³ French Imports to German Exports	German Imports from France	French Exports to Germany	Ratio French Exports to German Imports	German Exports to U.S.	U.S. Imports from Germany	Ratio³ U.S. Imports to German Exports	German Imports from U.S.	U.S. Exports to Germany	Ratio U.S. Exports to German Imports
1900												
1st	5			25				—		—	1	—
2nd	2	No	No	30	No	No	No	—	No	—	1	—
3rd	3	data	data	50	data	data	data	776	data	487	1,049	2.15
4th	6			25				—		—	2	—
1907												
1st	28	187	*6.68*	114			10	212	*21.20*		—	
2nd	349	219	.63	164	No	No	8	—	—	No	—	No
3rd	24	9	.37	334	data	data	5	2	.40	data	229	data
4th	80	174	2.17	29			187	259	1.37		309	
1928												
1st							4				4,425	
2nd	No	No	No	No	No	No	5	No	No	No	1,208	No
3rd	data	data	data	data	data	data	5	data	data	data	140	**data**
4th							5				140	
1935												
1st	5	14	*2.80*	—			—	3	—		36	
2nd	4	5	*1.25*	No	157	No	—	1	—	No	—	No
3rd	4	1	.25	data	259	data	5	—	—	data	—	data
4th	5	251	*50.20*		1,416		5	4	*.80*		—	

¹ A dash means negligible amounts. For sources and notes see Appendix.
² In each pair the statistics are those of the first country named.
³ Pairs of data with italicized ratios are excluded from the charts due to the smallness of the original quarterly figures or to the large size of the ratio.

CHART 1

Ratio and Magnitude of Gold Movement Statistics
of Pairs of Countries, Yearly Sums

POUNDS STERLING (THOUSANDS)

200,000 — ○ F 35 US

• United Kingdom
○ United States
▲ Germany

100,000 — UK 35 US ● ● F 35 UK

50,000 — ● UK 35 F

C 35 US ○ C 28 US
UK 28 F ● ○ ● UK 07 US

10,000 — G 07 UK ● US 28 UK
UK 07 F → ● ●
US 00 UK ● ● UK 28 US

F 07 UK ● ○ ○ ● UK 00 G
C 07 US → ○ US 28 C
US 00 F ○ US 07 F
US 07 C ○ ● UK 07 G

● G 00 UK ● UK 00 F
● F 00 UK C 35 UK
C 28 UK

○ C 00 US

1,000 — ○ ● UK 00 US
F 07 US ▲ US 00 G
○ US 00 C ○ US 07 UK

500 — ● G 35 UK ▲ G 07 F
○ F 00 US

G 07 US
▲

● US 35 UK

100 —
.00 .30 .60 .90 | 1.20 1.50 1.80 2.10 2.40
            1.00

RATIO

18

values of the transactions (using log-scale) are shown in 1000 pound sterling, the value taken from the first named country in each case.[8]

The *additional hypothesis* intimated in these charts is that when gold movements are large the ratio will improve,[9] perhaps because greater care in reporting is usually taken by both sides in such cases. Whatever the reason, the noted relativity seems to be borne out, first of all, by the yearly data. But even there it is noteworthy that the four largest values are all of the year 1935, and the really interesting gold standard period (pre-World War I) again fares badly. Even more noteworthy is the fact that the second largest value, involving not less than 85 million pound sterling of shipment between the United States and the United Kingdom was, according to the statistics of the latter, only 75 per cent of that which the United States reported for the same time—hardly a reliable piece of information! Physics would not have gotten very far with this sort of measurement.

(3) *Failure of the Statistics*:

The two charts of quarterly data (with the stated omissions of extreme and, indeed, quite absurd ratios) are far more revealing and hence more important. Even a casual inspection shows that the departures from unity are larger than anything admissible for reliable data. There are, it is true, especially for 1928 and 1935, a number of correspondences as good as can perhaps be expected —mostly involving the United Kingdom and Canada—but the scatter in the upper brackets is great and irregular, even for United

[8] Charts 1-3 show the frequency distributions of the ratios of Tables 1-2 in the following manner: To take an example: In Chart 1 UK 35 US at ● means that the statistics of the United Kingdom indicate a gold movement to the United States in 1935 (summed over the whole year). The dot shows (from the legend in Chart 1) that the United Kingdom Statistics were used to determine the ordinate, i.e. in this case the placement of the dot tells that the cumulative amount transacted was given as £85,883,000. The ratio at the bottom of the charts is taken from Table 1; that is, in the numerator are the data of the country whose statistics were not used for the determination of the ordinate (in this case the United States) and in the denominator the data of the other country (in this case the United Kingdom). In the case of UK 35 US the ratio was .75 as Table 1 shows and to this corresponds the abscissa of the dot ● on Chart 1. Correspondingly the other points on these charts have to be interpreted. In Charts 2 and 3 the quarterly sums are shown, *but no distinction* is made as to which quarter of the year is involved.

[9] The charts omit the extreme ratios altogether for reasons of difficulty in plotting, but they can be seen on Tables 1 and 2.

19

CHART 2
Ratio and Magnitude of Gold Movement Statistics
of Pairs of Countries, Quarterly Sums
Pre-World War I

20
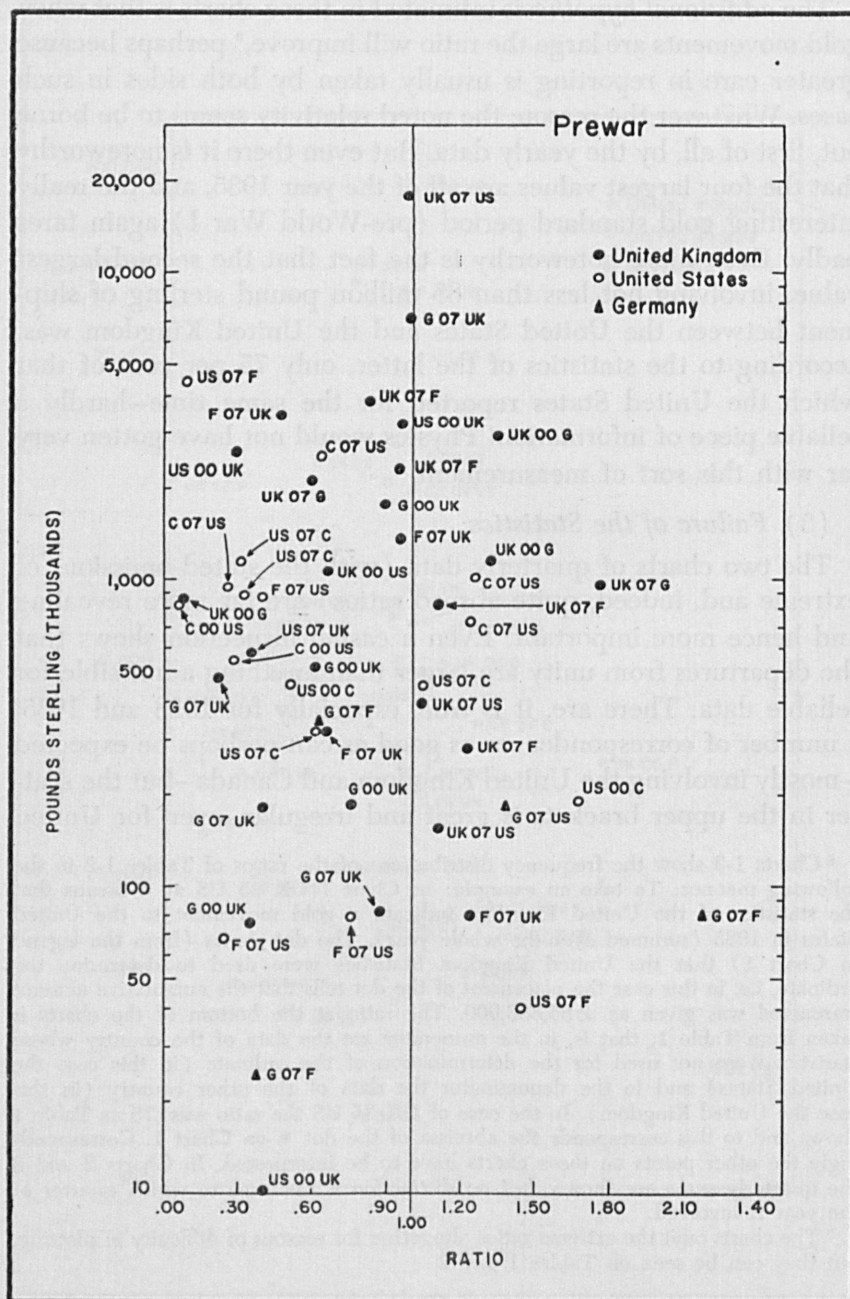
States and Canada, which is traffic across the border of neighbors.

The more interesting records are for 1900 and 1907, years of the classical gold standard. The result is catastrophic, especially since the largest ratios are omitted and, as noted in Tables 1 and 2, there are many instances (for quarterly and yearly sums) when one country reports a movement and the other does not,[1] so that a computation of ratios becomes meaningless. There is no recognizable tendency, no systematic bias of any sort. The statistics are simply contradictory. The quarterly sums for the period after World War I fare somewhat better (there are fewer entries) but are still useless for purposes of more intricate economic reasoning. Nonetheless, they were used in thousands of publications, scientific and political, in which all kinds of delicate "inferences" were drawn.
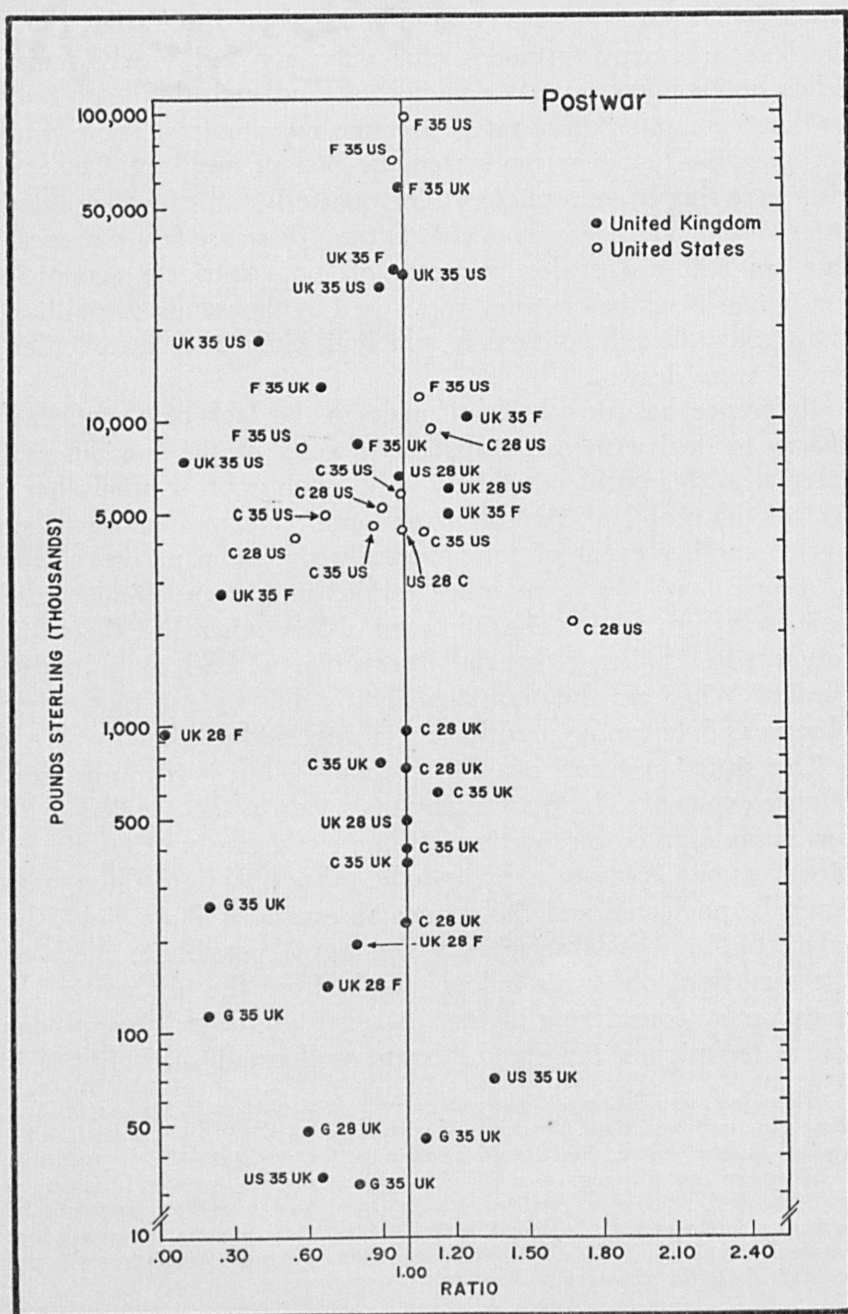
Reference has already been made to the lack of a *statistical theory* to deal with this situation. In spite of the fundamental interest in this problem, we can only touch upon it briefly here. This is the problem: Two observers note the same non-recurrent event; another event of fundamentally the same nature occurs after one month (or some other period) and observations of the new event are made. The observers differ; when are their differences so significant that the observations of both must be discarded? When are they not significantly different, so that either observer's data can be used for a well specified purpose?

This situation occurs often in economics, but it is rarely thoroughly explored. There are many instances in the applied field, for example, in crop reporting.[2] If one could be sure that of two observers one is trustworthy and the other not, their differences could be neglected and the former observations alone would be taken. In the physical sciences it is, as a rule, possible to check on the quality of observers independently from the observation in question by considering his results regarding other observations where techniques and standards are well established. But it is

[1] Here one must distinguish between (1) a lack of data and (2) a report that there were no transactions. We have, of course, excluded (1); if (2) prevails for a whole quarter or year, then the statement in the text applies.

[2] There are many examples in the physical sciences, too, such as observations of novae and supernovae, meteors, and extraordinary sound propagation; and numerous instances may be found in the military field and, in general, where witnesses are involved. All of this has been discussed at some length in my previously cited On the Accuracy of Economic Observations.

21

CHART 3

Ratio and Magnitude of Gold Movement Statistics
of Pairs of Countries, Quarterly Sums
Post-World War I

Postwar

POUNDS STERLING (THOUSANDS)

100,000 — ○ F 35 US

F 35 US ○

● F 35 UK

50,000 —

● United Kingdom
○ United States

UK 35 F
UK 35 US ● ● UK 35 US

UK 35 US ●

F 35 UK ● ○ F 35 US

10,000 — F 35 US ○ ● F 35 UK ○ ● UK 35 F
● UK 35 US ○ C 35 US ○ C 28 US
C 28 US → US 28 US
C 35 US → ○ ● ○ ● UK 28 US
5,000 — C 28 US ○ ○ ● UK 35 F
C 35 US ○ ○ C 35 US
● UK 35 F US 28 C

○ C 28 US

1,000 — ● UK 28 F ● C 28 UK
C 35 UK ● ● C 28 UK
● C 35 UK
UK 28 US ●
500 — ● C 35 UK
C 35 UK ●
● G 35 UK ● C 28 UK
● ← UK 28 F
● UK 28 F
● G 35 UK
100 —
● US 35 UK
50 — ● G 28 UK ● G 35 UK
US 35 UK ● ● G 35 UK
10 —

.00   .30   .60   .90   1.20   1.50   1.80   2.10   2.40
                    1.00
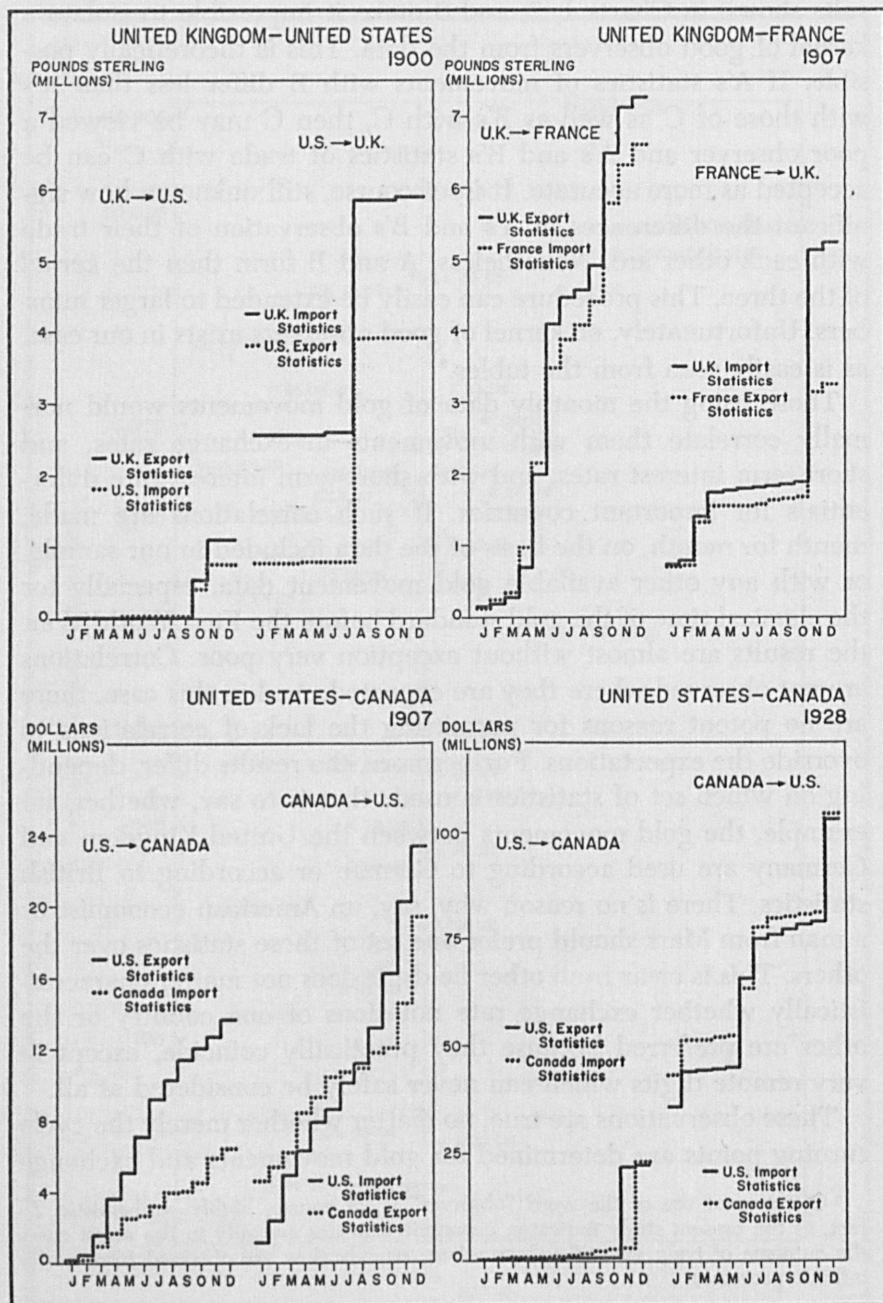   RATIO

22

not possible to say whether the British or the American foreign trade statistics are or were more accurate, or whether one was continuously more reliable than the other. Moreover, the poor results shown in Charts 1, 2, and 3 make it impossible to isolate a kernel of good observers from the data. This is theoretically possible. If A's statistics of movements with B differ less than A's with those of C as well as B's with C, then C may be viewed a poor observer and A's and B's statistics of trade with C can be accepted as more accurate. It is, of course, still unknown how significant the differences in A's and B's observation of their trade with each other are. Nonetheless, A and B form then the *kernel* of the three. This procedure can easily be extended to larger numbers. Unfortunately, no kernel of good observers exists in our case, as is easily seen from the tables.[3]

Those using the monthly data of gold movements would normally correlate them with movements in exchange rates, and short-term interest rates, and with short-term interest rate differentials for important countries. If such correlations are made, month for month, on the basis of the data included in our sample, or with any other available gold movement data, especially for the classical time of the gold standard before the First World War, the results are almost without exception very poor. Correlations are not observed where they are expected. And in this case, there are no potent reasons for permitting the lack of correlations to override the expectations. Furthermore, the results differ, depending on which set of statistics is used: that is to say, whether, for example, the gold movements between the United Kingdom and Germany are used according to German or according to British statistics. There is no reason why, say, an American economist or a man from Mars should prefer one set of these statistics over the others. This is clear from other fields; it does not matter characteristically whether exchange rate notations of one country or the other are preferred because they practically coincide, except in very remote digits which can never safely be considered at all.

These observations are true, no matter whether merely the cycle turning points are determined for gold movements and exchange

[3] Our present use of the word "observer" is, of course, highly euphemistic. In fact, as the present study indicates, economic statistics are only in the rarest cases the outcome of truly scientific observations; usually they are obtained from vastly inferior processes.

CHART 4

Cumulative Monthly Gold Movement Statistics for
Pairs of Countries from Respective Sources
1900, 1907, 1928

UNITED KINGDOM—UNITED STATES 1900

POUNDS STERLING (MILLIONS)

U.S.→U.K.

U.K.→U.S.

— U.K. Import Statistics
••• U.S. Export Statistics

— U.K. Export Statistics
••• U.S. Import Statistics

J F M A M J J A S O N D   J F M A M J J A S O N D

UNITED KINGDOM—FRANCE 1907

POUNDS STERLING (MILLIONS)

U.K.→FRANCE

FRANCE→U.K.

— U.K. Export Statistics
••• France Import Statistics

— U.K. Import Statistics
••• France Export Statistics

J F M A M J J A S O N D   J F M A M J J A S O N D

UNITED STATES—CANADA 1907

DOLLARS (MILLIONS)

CANADA→U.S.

U.S.→CANADA

— U.S. Export Statistics
••• Canada Import Statistics

— U.S. Import Statistics
••• Canada Export Statistics

J F M A M J J A S O N D   J F M A M J J A S O N D

UNITED STATES—CANADA 1928

DOLLARS (MILLIONS)

CANADA→U.S.

U.S.→CANADA

— U.S. Export Statistics
••• Canada Import Statistics

— U.S. Import Statistics
••• Canada Export Statistics

J F M A M J J A S O N D   J F M A M J J A S O N D

24

rates or the other series referred to, or correlations are time series correlations involving all consecutive values. Since this is a question of application of series and is not directly connected with the test of the validity of the statistics, the material on which these remarks are based is not published in this paper.

Further information about the inner nature of the discrepancies is highlighted in Chart 4 which shows, for some cases—randomly chosen—*cumulative sums of monthly data* from two sources. It is interesting that the statistics involving Canada and the United States are among the most unreliable; this is especially the case because of the shift of the discrepancy first from one, then to the other. We have no way of telling whether the Canadian or the United States statistics, or both, have to be rejected. The differences are particularly notable for the United States exports to Canada for 1907. But the other pairs of countries do not fare much better. They have one thing in common: when differences exist, they very frequently persist for many months in succession and are often not removed. This means that there are, apparently, at the best no self-correcting forces at work making themselves felt swiftly.[4] In general: even if the direction of movement should be the same, the great *absolute* differences make it impossible to use these statistics for purposes of determining for reasonably short intervals the amounts of gold shipped and, possibly, of capital transfers. They can only be used for the crudest kinds of estimates and can play no major role in the theory of international trade.

(4) *Further Explorations*:

Since most writers on international trade are particularly concerned with *net movements*, we insert two Charts—again randomly chosen—showing the net movement of gold. Chart 5 shows the movement from the United Kingdom to Germany for the period from 1900-1912 and 1931-1937 on a *monthly* basis and Chart 6 shows, for the period from 1925 to 1937, the net gold movements from the United Kingdom to the United States. In each case the statistics are made up from the two possible sources: in Chart 5 from United Kingdom figures and German figures, and

---

[4] There is a *lag* in several cases and there may be transit movements, as already mentioned. Elaborate statistical operations on these data may be indicated in order to discover lags, especially when more than two countries are simultaneously considered.
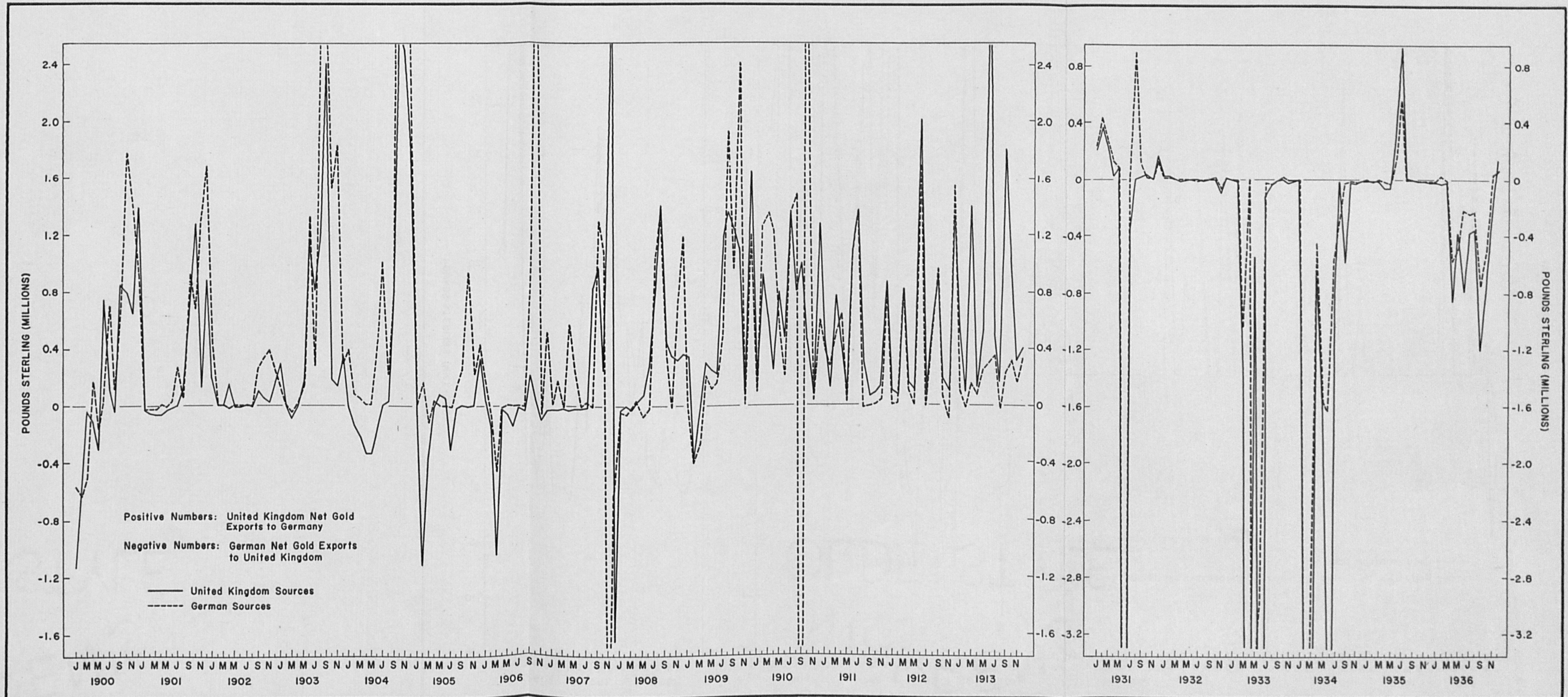
in Chart 6 from British figures and United States figures. In both graphs, the two lines would coincide if the data were accurate, but this is the case only in rare instances. Anyone interested in the cyclical aspects of gold movements would be in a considerable quandary in determining where the cycles are, what the amplitudes are, and into which particular month the peaks and troughs of the cycles fall. Sometimes there is for the two sources a clearly indicated, but inconsistent, lag—often considerable, comprising four to five months; sometimes movements occur for which there is no record in the corresponding other statistics; and sometimes there is good agreement. We forego any detailed interpretation of the charts.

The confusion of the statistics as between the two European countries is much greater for the time before the First World War than it is in the early thirties, but it is worth recalling that this latter period coincides with a time in which exchange control was applied quite strictly by Germany and so it would be expected that more accurate records would be kept. On the other hand, Chart 6 for the United Kingdom and the United States shows relatively few conflicts in the data in the first part of the period, but considerable deterioration in the second part. Even the last years show a rather erratic behavior and the situation is one which should cause business cycle applications of these statistics to be viewed with grave concern. The frequently stated opinion that statistics of recent times tend to be better than those of earlier times is not necessarily true.

It is also interesting to note that in the years preceding the outbreak of World War II the gold transactions between the United Kingdom and the United States were many times as large as those between the United Kingdom and Germany. The magnitude of the operations is not necessarily a guarantee that the operations are more carefully recorded, our remarks on page 19 notwithstanding.[5]

[5] A somewhat laborious extension of our procedures suggests itself: instead of forming cumulative sums (gross or net) merely for significant pairs of countries, all countries, or at least the most important, could be taken together and an overall total of imports and of exports could be formed. It might be hoped that in this way all major discrepancies would disappear. Attempts made in this direction have not succeeded—as can be gleaned from the data presented in this study —and even if a favorable result should be obtained, it is difficult to see how the all-important country to country data could be corrected.

CHART 5

Comparison of Statistics of Monthly Net Gold Movements Between United Kingdom and Germany, 1900-1913, 1931-1936

Positive Numbers: United Kingdom Net Gold
            Exports to Germany

Negative Numbers: German Net Gold Exports
            to United Kingdom

——— United Kingdom Sources
- - - - German Sources

CHART 6

Comparison of Statistics of Monthly Net Gold Movements Between United States and United Kingdom, 1925-1938

POUNDS STERLING (MILLIONS)

POUNDS STERLING (MILLIONS)

——— United Kingdom Sources
------ United States Sources

Positive Numbers: United Kingdom Net
Gold Exports to United
States

Negative Numbers: United States Net
Gold Export to United
Kingdom

J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N J M M J S N

1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935

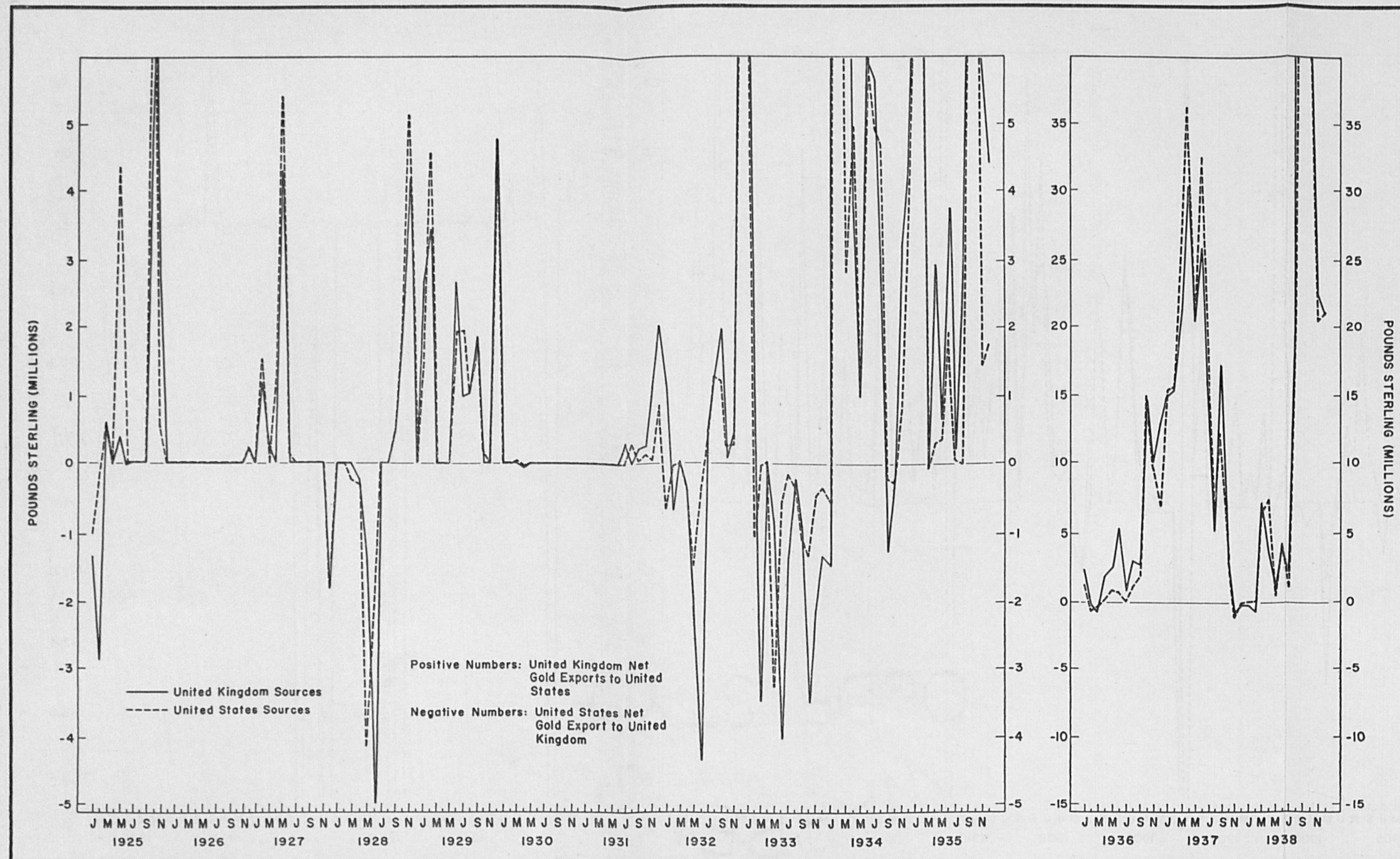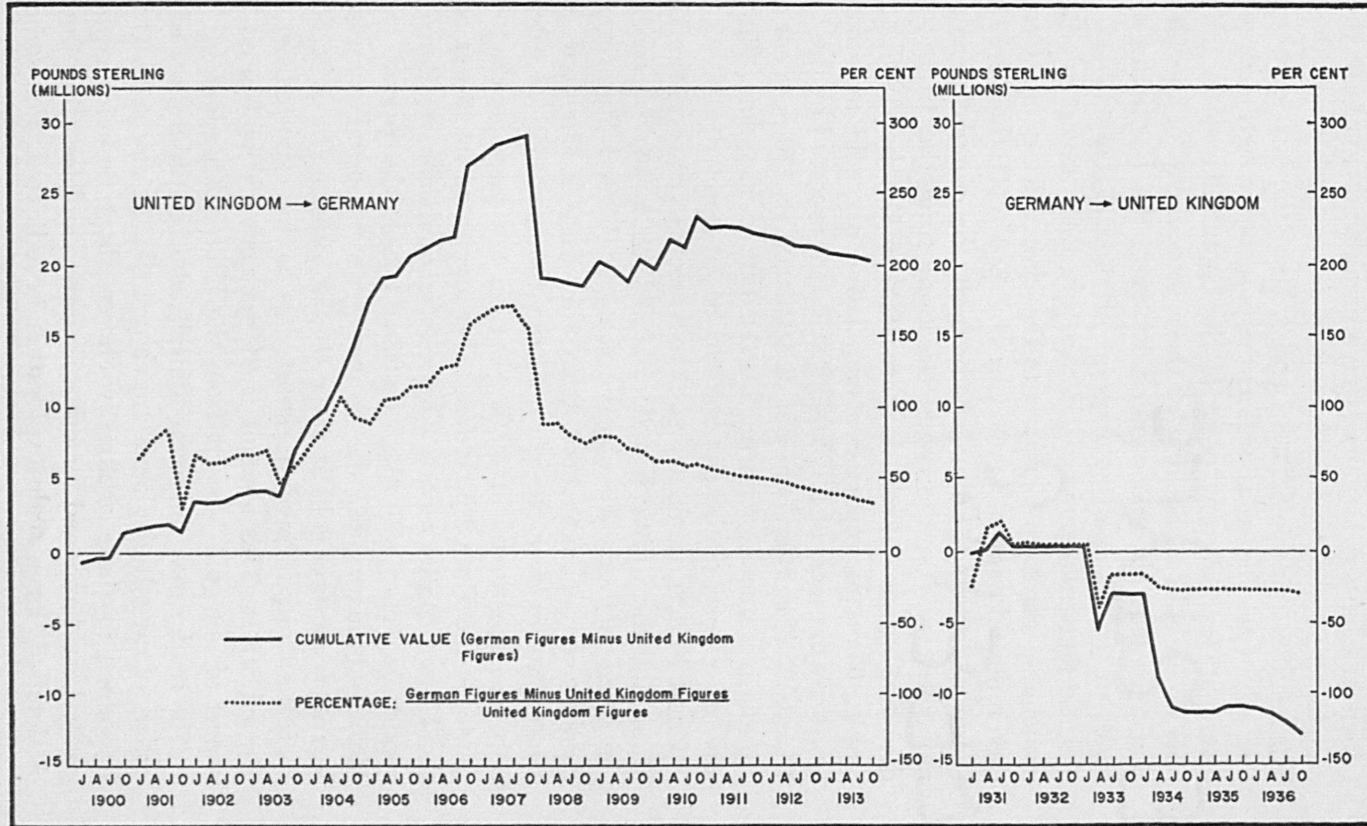J M M J S N J M M J S N J M M J S N

1936 1937 1938

CHART 7

Cumulative Discrepancies in Reported Net Gold Movements Between United Kingdom and Germany, 1900-1913, 1931-1936
(Cumulative Values and Percentages)

POUNDS STERLING (MILLIONS) PER CENT

UNITED KINGDOM → GERMANY

GERMANY → UNITED KINGDOM

CUMULATIVE VALUE (German Figures Minus United Kingdom Figures)

PERCENTAGE: German Figures Minus United Kingdom Figures / United Kingdom Figures

J A J O J A J O J A J O J A J O J A J O J A J O J A J O J A J O J A J O J A J O J A J O J A J O J A J O J A J O
1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913

J A J O J A J O J A J O J A J O J A J O J A J O
1931 1932 1933 1934 1935 1936

In order to facilitate the evaluation of the information contained in Charts 5 and 6 we have cumulated quarterly the discrepancies in the reported net gold movements for both cases. The result is found in Charts 7 and 8. There the quarterly cumulative values are shown as well as the percentage of the cumulative discrepancies. Any one point of the cumulative series indicates how much gold is, because of the discrepancies of the information, cumulatively unaccounted for from the beginning of this series up to that particular point of time. Taking Chart 7, we observe that from January 1900 to July 1900 the cumulative net movement was from Germany to the United Kingdom. Germany then reported a net movement less by £ 310,000 than the United Kingdom reported. From July 1900 on, the cumulative net gold movement was in the other direction: from the United Kingdom to Germany. From January 1900 to October 1907, for example, Germany reported a cumulative net gold movement to Germany in excess of £ 29,000,000 to that reported by the United Kingdom. This discrepancy of £ 29,000,000 is 156 per cent of the cumulative amount reported by the United Kingdom. Similarly, any other parts of the chart are to be interpreted.

Chart 8, describing net gold movements between the United States and the United Kingdom 1925-1938, is made according to the same method. We observe that the United States-reported cumulative net movement from the United Kingdom to the United States over the period from January 1925 to July 1932 only was greater by £ 9,100,000 than the corresponding report made by the United Kingdom. This amounted to 26 per cent of the United States report. Curiously enough the relation of the reports switches into the opposite: taking the period January 1925 to January 1937 the United States reported cumulatively *less* net receipts of gold from the United Kingdom by £ 52,500,000 than the United Kingdom reported as having exported net to the United States. This amount is 21 per cent of the United States report of net gold imports from the United Kingdom.

The reader may judge for himself whether differences of these orders of magnitudes mean much or little. At any rate, what has happened in those arbitrarily chosen short time intervals may easily have happened for other countries and for other intervals. A detailed, large scale exploration is indicated. For some pur-

poses the discrepancies may not matter, but if they do not, then it is because economic reasoning in these fields cannot be as sharp as it is often assumed to be.

CHART 8

Cumulative Discrepancies in Reported Net Gold Movements Between
United States and United Kingdom, 1925-1938
(Cumulative Values and Percentages)
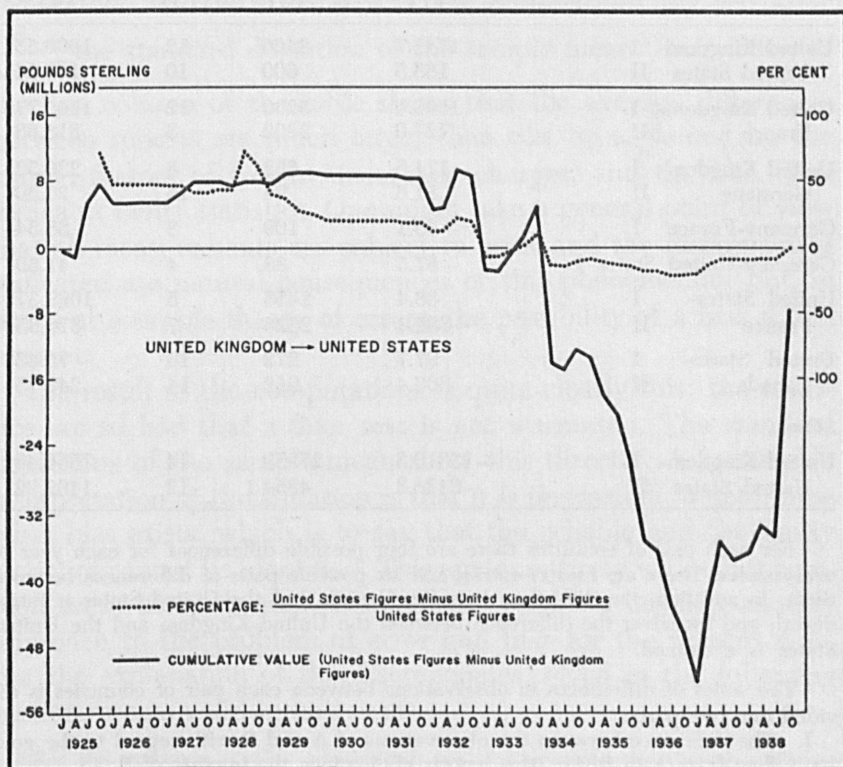
The arguments used so far in evaluating the statistics were mostly qualitative because the lack of statistical theory makes the desired quantitative expression at present impossible. But in order at least to describe the data statistically along conventional lines, the significance of algebraic differences of the quarterly statistics of gold shipments between pairs of countries is shown in Table 3. The data used here were taken from Table 2 and include quarterly statistics of gold movements between all possible pairs, covering trade of the United States, Germany, France,

TABLE 3

TEST OF QUARTERLY GOLD STATISTICS
(1900, 1907, 1928, 1935)
Algebraic Differences[1] between Pairs of
Countries and a Test of Their Significance.

Pairs of Countries[2]		Average Differences[3] (1)	Standard Deviation[4] (2)	n (3)	Standard Deviation of Sample Mean $\frac{\sigma}{\sqrt{n}}$ (4)
United Kingdom-	I	1741.7	3466	12	1000.55
United States	II	168.5	600	10	189.74
United Kingdom-	I	1696.6	5230	12	1509.77
France	II	727.0	2302	8	813.88
United Kingdom-	I	−171.5	652	8	230.52
Germany	II	81.6	94	16	23.50
Germany-France	I	−45.1	109	8	38.54
Germany-United States	I	−67.5	83	4	41.50
United States-	I	38.4	2458	5	1099.37
France	II	832.4	2326	7	876.85
United States-	I	97.7	272	13	75.43
Canada	II	303.4	912	14	243.75
Silver					
United Kingdom-	I	−13810.3	27959	14	7556.48
United States	II	−6426.8	4854.1	12	1402.92

[1] For each pair of countries there are four possible differences for each year of every series. There are four countries and six possible pairs of differences between them. In addition, the difference between Canada and the United States is introduced; and for silver the difference between the United Kingdom and the United States is examined.

[2] The series of differences in observations between each pair of countries is divided into two parts:
I. The difference between the observations of A and B with regard to the gold flow from A to B; i.e., the exports of A minus the imports of B.
II. The difference between the observations of A and B with regard to the gold flow from B to A; i.e., the imports of A minus the exports of B.

[3] Statistics of first country minus those of second. These algebraic differences are then summed and divided by their total number. In some pairs of countries, data are not reported for all years. This diminishes the number of differences to less than 16 (quarterly figures taken for four years) since years for which no data are reported are excluded.

[4] The computation of the standard deviation of the sample mean is sometimes conventionally restricted to n ≧ 10. We nevertheless include the values for smaller n since the tendency of the different computations appears to agree in all relevant aspects.

30

Great Britain, and Canada for the years 1900, 1907, 1928, and 1935. Silver transactions between the United States and the United Kingdom are also included in order to be able to compare the two precious metals in at least one relationship.

In Table 3, the first column shows the average differences with regard to the two directions of the gold flow between two countries. The second column shows the standard deviations of the population, the third the size of the sample, and the last gives $\frac{\sigma}{\sqrt{n}}$, the standard deviation of the sample mean.[6] Inspection of the first column of the table shows that the average differences between reports are much larger than can be explained by the known factors of freight, insurance charges, and the like. This indicates faulty statistics. One might take a general point of view that all measurements are subject to error and the discrepancies indicated are natural consequences of this phenomenon; but, in terms of a simple theory of errors, the possibility of a bias is still present.

The result of the computations is quite clearly this: the statistics are so bad that a finer test is not warranted. The standard deviations of the sample mean show this directly. A reasonable interpretation of the situation is that it is impossible to determine what bias exists, which is to say that the reliable and the faulty statistics cannot be identified. This agrees entirely with what common sense has indicated. It may be possible to develop another approach to the problem of error and bias for the present case, but the explanation of the discrepancies, given in the following section, seems valid in any case.

[6] For the rationale of the calculations, see, for example, H. Cramèr: *Mathematical Methods of Statistics*, Princeton, 1946, p. 345, or A. M. Mood: *Introduction to the Theory of Statistics*, New York, 1950, p. 133.

31

IV. Explanation of the Differences
and Further Difficulties

Where do the differences come from? It would be intriguing and instructive to explore this matter in full detail, but space forbids more than a brief enumeration of the chief factors.[7] One factor has already been mentioned: economic statistics are seldom the result of scientific observations in the sense of the statistics used in the sciences. They often are unwillingly produced by-products of other activities. There is also, of course, the slight difference in the position of the "observers"; this consists, for gold, mostly in the time interval that may occur. One must not make too much of this factor for it is negligible for bordering countries where gold was shipped on fast trains; and even for overseas shipments a constant lag could only be explained if, say, English exports to the United States occurred only, or chiefly, at the end of the month and thus were always registered in United States statistics for the next month—an absurd idea. When "lags" appear, as we have noted they often do, they are not stable but sometimes go in one direction, sometimes in the other. They must therefore have other reasons. Besides the genuine time lag due to transportation between distant countries, there clearly are others that are (a) spurious and due to inaccuracy and (b) due to transit.[8] Most likely, (a) holds more frequently.

All differences due to prices, freights, insurance, duty (zero!), exchange rates (excepting Germany which, in 1935, had a variety of exchange rates) can be neglected. They are more or less known and trifling.

So there remain the "true" causes. We leave open, whether the statistics were often, occasionally, or never deliberately falsified. We do not know this, but since we are dealing with gold, about transactions in which a great deal of secrecy often prevailed, it

[7] A full investigation is called for of the other figures of international trade as well. This is not a matter solely for academic economists, but is the duty of those collecting the original statistics and of the government offices using them.

[8] Amounts in transit could perhaps be discovered by means of the rather laborious computations mentioned above, involving all countries between which gold movements occur—if the figures were worth it.

is not impossible that such manipulation occurred. Inaccuracy connected with privilege is frequent; central banks, and sometimes even private banks, often could remove gold from ships in a great hurry and the statistical records may then have suffered even unintentionally. It is known, for example, that gold was frequently put on a barge at the Ambrose light ship before the steamer arrived in New York and taken directly to the importing bank's office. It is also to be noted that the high value of the metal is not a sufficient cause for accurate entries, especially when the customs officials had no interest whatsoever in the valuation of gold.

There are possible explanations arising from differences in *classification*. Although "gold" is easily determined as such, gold and silver were often thrown together in the same cases and shipments, and frequently there were in such instances false declarations on the part of the shippers. Freight and insurance were cheaper on a "mixture" of gold and silver and it was easy even to add a piece of silver to a gold bar. In one country the mixture might leave as silver, and arrive in the other as gold and vice versa.[9] Besides savings in freight and insurance there was often also the desire for secrecy as to shipments when private and commercial banks were involved and were assuming the risks.

Coverage of all gold movements by the statistics was unquestionably very incomplete. We have already referred to secrecy, removal of gold at arrival in port before customs, and so on. There are, furthermore, the not negligible gold movements in the form of gold carried by travelers during the periods of gold coin circulation. While the amounts were large, they were neglected in *all* the statistics and would therefore, even if included, not have closed the differences in reporting shown by this study.[1] Of course, if some countries included this item and others did not, the con-

[9] Table 3 includes a separate column for *silver*. The results are not better for this metal as further checks have revealed. Nor does a total of gold and silver give better results, which shows the absence of any self-correcting elements in the statistics. Silver is, of course, more in the nature of a commercial commodity than gold. This gives an idea what to expect from tests of foreign trade statistics. See my book, *On the Accuracy of Economic Observations*, where an attempt is made in that direction.

[1] There is occasional mention in the literature of this gap in the statistics; but it is clear that it is a peripheral phenomenon as compared with the other deficiencies we have noted.

fusion is compounded. The discovery of *origin and destination* of gold shipments probably offers as great an objective difficulty as with all other commodities.[2] Gold arriving in B from A may be destined for C but packed in cases of D. In different countries different practices exist as to how to enter such amounts (of "specie") into the statistics. For example, transit figures are not to be had with accuracy although transit played a great role in some trade, especially that involving the United Kingdom and Canada. Sometimes gold would travel through several countries in transit status; this was frequently the case for European countries in particular. Gold already at sea would often be rerouted due to a change in exchange rates, but the export statistics of the sending country, even if they had been correct, would not and could not be changed correspondingly. Clearly, the lumping together of their own gold movements with gold in transit produces an aggregate that cannot be used in analysis, even if it should be correct as to amount.

Among a host of other *further difficulties* one—*earmarking of gold*—needs attention for difficulties arise even when the statistics are accurate, which may or may not have been the rule. The question of accuracy must be asked of the central banks in this case, rather than of the authorities responsible for foreign trade statistics.

The beginning of the practice of earmarking, the sale of gold in country A to country B without actual shipment but release of the quantity from A's and its possible addition to B's central bank reserve, is obscure. It dates back to at least 1900 at which time *The Economist* (London), in criticizing the *Statist*, "corrected" British statistics for certain movements of Indian gold,[3] though it is difficult to see what information the former periodical was using. Earmarking became widespread from the late 1920's on, for it was one of the general aims of international cooperation of central banks to eliminate physical gold movements as far as possible.

[2] The efforts of the various international statistical organizations to obtain conventions, to regulate terminology, etc. are great, but of fairly recent origin. The discussions and proposals thereon give a vivid idea of the obstacles to be overcome. See especially the various reports of the Committee of Statistical Experts, League of Nations, and, more recently, the important work by Allen and Ely, mentioned in footnote 6, p. 6.

[3] India used to keep a considerable portion of her gold reserve in London.

34

Various international institutions served this purpose. It is probably safe to say that earmarking assumed ever increasing importance.

Now it is clear that it is the book transfer of the monetary reserve that matters for central bank operations, not the information of a later physical shipment, which might well be from A to D—with which country A had no dealings at all—B having sold the earmarked gold from A to C and C perhaps to D and actually shipping it there. In the first case, the point of time of the shipment is wrong from the economic point of view, in the second the—correct!—movement information is completely meaningless so far as A is concerned. One full shipment can easily be recorded as several transactions. All this can be still further complicated when transit intervenes. Central banks do not, as a rule, say for whom they hold earmarked gold and sometimes will not say whether they hold it at all. There may thus be many changes in the composition of ownership of an otherwise unchanged total. This makes a correction of the foreign trade figures impossible. At best, statements about earmarking serve to indicate changes in the aggregate gold stock of the central bank that either sells or buys such gold. The Bank of England, when releasing gold from earmark, always has refused to say for whom the release was made. The amounts involved in earmarking became larger and larger as the practice grew.

The *consequences of the test*, as far as the foreign trade statistics of gold are concerned, were noted above and need not be repeated.[4] But there remains the question of *substitutes* from which pairwise gold movements might be deduced, from which, in turn, inferences about capital movements could be made. This chain is long, and there are no substitutes. It will be recalled that statistics of gold movements are themselves only somewhat doubtful and incomplete substitutes for those of capital movements. There remain only the published central bank figures, but they do not inform us about pairwise movements, nor can it be expected that pairwise gold transactions could still be uncovered from the

[4] The inaccuracy of the data gives an idea of the difficulty of obtaining numerical values in economic analysis, even when the stochastic character of the magnitudes entering into the fundamental equations is taken expressly into account. However, in the present case the difficulty is that one cannot determine clearly what "accuracy" means and, consequently, cannot measure deviations from it.

unpublished records. Furthermore, even central bank information would only cover a part of the gold movements since often the larger part came from the open market, largely the commercial and private banks. It is impossible ever to reconstruct those transactions from any sources. For general economic histories, especially those dealing with large crises, a combination of foreign trade and central bank statistics may roughly suffice; but for purposes of finer detail of economic theory such is not the case. Moreover, a careful test of the accuracy of central bank figures would also be necessary.[5] They are sometimes deliberately falsified by combining data on gold and silver and by reporting bullion gold or foreign coins under "other assets," a common practice in the period between the World Wars. Finally, there is the general secrecy regarding operations of exchange stabilization funds whose gold operations often overshadow those of the central banks. The irregular accounting practices noted here were also common in the nineteenth century but economists have paid only scant attention to them.[6]

Our discussion of gold movements has been in terms of *pairwise trade* between countries. This is unquestionably the basic information needed when exchange rates are considered and when general trade with another country is studied. One may also be interested in the aggregate position of the balance of payments of a country. In that case, the total net gold movement is of interest. It is clear, however, that this figure is ultimately based upon those pertaining to the pairwise relationships which we have found to be very defective. If it could be proved that the aggregation has some self-correcting feature, such that the errors in the individual components of the aggregate tend to cancel out, the total would be more valuable than the parts. Yet there appears to be nothing in the general picture that would allow us to draw this comforting conclusion. No self-correction can be observed.

[5] The author also attempted to break up gold transactions of central banks according to countries in order that such time series might be used as substitutes for those derived from foreign trade statistics. This idea was kindly tested for me for 1898 and 1907 by the Bank of England, but the results were entirely negative. They disclosed that private gold shipments were very large if the customs' figures can be accepted. However, even if such series were constructed they would not form the desired substitute.

[6] In spite of their elaborate work, Palgrave and others have not asked enough questions about the meaning and background of some of their figures.

The net gold position of a country is thus only inadequately known. It is, however, theoretically possible to check the movement data from the changes in the position of the gold stocks in a country. This is not too difficult in those countries where a strict control exists—at least as far as monetary gold is concerned. It is a virtual impossibility for the period of the classical gold standard where the gold and gold coin could be in anybody's possession and where neither commercial banks nor the central bank made correct and complete disclosures of their assets, not to mention the movements in and out of industrial gold.

V. Summary and Conclusions

Our findings can perhaps be summarized in a few points:

a) If the statistics agree for both sets (preferably according to a powerful test of significance for which, however, a theory is lacking), they may still both be wrong. This cannot be decided, as in most other cases. Either statistics may then be used.

b) If the available data disagree, the question is whether the differences are significant after giving full consideration to the reasons discussed above as to why they might differ. Unfortunately, no significance test is available that is specifically suited to the situation encountered here.

c) There is, at present, no objective way of choosing between the various ultimate sources. The insistence of an American investigator on American and of a British investigator on British sources—each decision leading to different results and interpretations—cannot be overcome by any scientific arguments.

d) If, in the light of common-sense, the statistics are significantly different, there seems to be no alternative but to reject them all, along with all the detailed series derived from them. All the other consequences of such a step must, of course, also be taken, including the doubts that must arise as to the value of much of past and present literature, including not only the writings using these statistics explicitly but also those concerned with the pure theory of international trade to the extent that they are based, implicitly or explicitly, on them.

e) If there is a significant common-sense difference in the data and both sets are used to corroborate the theory or to give new correlations with other time series (which turn out to be high), and one set agrees well with the theory of international trade or some other *prima facie* plausible model, the statistics, and thus their use, should still be rejected. The theory with which they may happen to agree is not strong enough, for it usually has no firm and clear other empirical foundation, where the data have been checked up to the present standards and are beyond doubt of a superior quality than those for gold. These foundations of the theories are in most cases too weak to allow the overriding of the other non-corresponding set of data.

38

Therefore, when a significant difference appears, work based upon either set should not be undertaken—quite aside from the improbability that at least one set of data should corroborate the theory.

f) The statistics might be considered adequate for broader corroborations of the general theory of international trade having to do mostly with such long-range phenomena as general price tendencies and the overall movement of gold,[7] but not for the study of cyclical behavior or, for example, for forming conclusions as to whether gold moves before or after the proceeds of an international loan are used by the borrowing country. The first kind of problem may not require more than coarse data for the opening of the subject for a more detailed study, while the second cannot be undertaken unless the data are as fine and as accurate as the mechanisms they are supposed to corroborate.

It is clear that the dropping of these statistics is a serious matter for the theory of international trade. It is hard to see what the other empirical foundations are besides foreign trade figures, exchange rates, and price statistics. Of these the first are, if possible, less reliable than the statistics of gold movements. If one part of the theory rests—if it rests on anything at all—on such doubtful material, then suspicion rises against its other parts; but we shall not dwell upon this matter.[8]

The regrettable lack of trustworthy gold statistics deprives the study of international trade and business cycles of one of its most important means of analysis. The effects are felt not only in the inability to make reliable correlations between gold movements and other time series,[9] but also in respect to further exploitation of statistics which make essential use of the information on gold movements. This matter is particularly relevant with respect to balance of payment statistics which already abound with esti-

[7] For example, nobody will deny that huge amounts of gold flowed to the United States after 1933, a fact for which there is also other corroboration.

[8] There is, in particular, the notion of the "terms of trade" and its statistical measurements, quite apart from the underlying logic. Comparative costs are often weakly put on ancient labor value ideas; and the belief, that when countries are trading with each other as units, one has to solve ordinary maximum problems cannot any longer be maintained now that the theory of games has shown that in such instances no ordinary maximum problems exist.

[9] When such correlations have been made, they failed to show any relations that could be interpreted with our present knowledge.

mates covering large percentages of the total balances. We are also unable to rely as much as we would otherwise wish on the conclusions of some of the best known standard works in the field of international economics, whether they belong to the classical period or are of more recent date.

In conclusion, the question arises whether gold movement statistics are better or worse than others, especially in the field of international commodity trade. My experience with the latter is that many trade data are worse (in terms of the simple criteria used in this study).

In this study only one single component of foreign trade and foreign payment statistics has been analyzed. Writers on foreign trade will have to assume the burden of proof that the figures on commodity movements are good enough to warrant the manipulation and the reasoning to which they are customarily subject.

What shall be done about this situation? It requires a great deal of strength not to use ready-made and commonly available statistics, dealing with important events, when there are no substitutes. But great as the temptation may be to continue as if everything were in tolerable shape, a greater service is rendered to economics by insisting that a good theory cannot be built on shaky data.

Appendix

SOURCES AND NOTES FOR TABLES 1 AND 2

Description of Material Used in Comparing Gold Movements
of United Kingdom, France, Germany and United States
for 1900, 1907, 1928 and 1935
(Tables 1-2)

United Kingdom

Source:

Board of Trade, *Trade and Navigation, Accounts relating to Trade and Navigation of the United Kingdom* (in House of Commons, Sessional Papers).

Unit:

Pound sterling.

Notes:

Given as gold bullion and specie for the United Kingdom. In making comparisons, the pound sterling is used as base.

France

Source:

Direction Générale des Douanes; *Statistique mensuelle du Commerce extérieur.*

Unit:

1900—hectogram
1907-1928—kilogram
1935—francs

Notes:

Cumulative series for gold given in bullion and coin. Exports of gold in 1928 have no breakdown by countries and imports have the very large figure for "all other countries" of 421,765 kilogram. Monthly figures for 1900 are incomplete. Quantity converted into value at rate of .32258 gr. per franc 1900, 1907, January-June 1928, and rate of .0655 gr. per franc July-December 1928. In comparisons with the United Kingdom, the franc data are converted into pound sterling at par for 1900 and 1907 and at the annual average of exchange rate for 1935 and 1928. A correction was made in the November 1935 imports without revising preceding months. This correction made the cumulative figure for November smaller than the preceding month. Referred to in French source as Great Britain.

In comparison with Germany, the franc data were converted into mark and reichsmark at par in 1900 and 1907 and at the annual average of exchange rate in 1935.

In making comparison with the United States, the franc data were converted into the dollar at par in 1900 and 1907 and at the annual average of exchange rate in 1935.

Germany

Source:

Statistisches Reichsamt; *Monatliche Nachweise über den auswärtigen Handel Deutschlands.*

Unit:

Kilogram

Notes:

In 1900 data are given as gold coin, bars, and plate. In addition, there are

41

figures for gold including other metals. These are insignificant except for imports of 12,800 kilogram of gold and platinum ore from Great Britain and waste from gold and silver work imported from United States—111,400 kilogram. After 1900 no groupings of gold with other metals were made. With the exception of 1900, the price of the type of gold was given and from this the value was computed. Data for 1907, 1928, 1935 here given as gold coin (foreign and German), and gold leaf, and bullion.

In making comparisons with the United Kingdom, mark and reichsmark were converted into pound sterling at par in 1900, in 1907 and 1928 and at the annual average of the exchange rate for 1935. Referred to in German source as Great Britain.

In making comparison with France, mark and reichsmark were used as base, as was also done in the case of the United States.

The annual total of United States imports from Germany in 1900 given in the source does not equal the monthly total.

United States

Source:

1900–07, Bureau of Statistics, Department of Commerce and Labor; *Monthly Summary of Commerce and Finance.*

1928–35, Bureau of Foreign and Domestic Commerce, Department of Commerce, *Monthly Summary of Foreign Commerce.*

Unit:

Dollar

Notes:

The 1928-35 figures agree with those published in the *Federal Reserve Bulletin.* The source does not specify whether the figures for 1935 are on a devalued dollar or old dollar basis. No distinction is made between coin and bullion.

The dollar was converted into pound sterling at par in 1900, 1907 and 1928, and at the annual average of exchange rate for 1935. The annual total of imports for 1907 does not equal the monthly total. The latter figure is used. In the *Federal Reserve Bulletin* the figures are classified as those for England and in the summary as those for United Kingdom.

In making comparison with France, the dollar is used as base.

The dollar is converted into the mark at par in 1900, 1907 and 1928, and at the annual average exchange rate for 1935 in making comparison with German statistics.

Canada

Source:

1900–07, Census and Statistics Office, *Trade of Canada*, monthly report.

1928–35, Dominion Bureau of Statistics, External Trade Branch, *Trade of Canada*, quarterly report.

Unit:

Canadian dollar

Notes:

Reports start in July 1900 and data are given as gold coin and bullion. It is not specified whether this includes some silver.

In making comparisons with United Kingdom and United States data, the Canadian dollars are converted into pound sterling at par.

RECENT PUBLICATIONS OF
INTERNATIONAL FINANCE SECTION

Survey of United States International Finance.
  By the International Finance Section staff.

1. Volume covering 1949	$1.75
2. Volume covering 1950	$2.25
3. Volume covering 1951	$2.25
4. Volume covering 1952	$2.75
5. Volume covering 1953	$2.75

Order from any bookseller or from PRINCETON UNIVERSITY PRESS

PRINCETON STUDIES IN INTERNATIONAL FINANCE

1, 2, and 3   Out of print
4. Postwar Bilateral Payments Agreements.
  By Merlyn N. Trued and Raymond F. Mikesell.
  (April 1955) $0.25

Order from International Finance Section. Free of charge to non-United States residents.

The International Finance Section also publishes from time to time papers in the two series: ESSAYS IN INTERNATIONAL FINANCE and SPECIAL PAPERS IN INTERNATIONAL ECONOMICS. These are distributed without charge by the Section to interested persons. Copies may be obtained by addressing requests directly to the International Finance Section, Princeton University. Standing requests to receive new numbers as they are published will be honored. Only the following are still in print.

ESSAYS IN INTERNATIONAL FINANCE

Problems of the Sterling Area, with special reference to Australia. By Sir Douglas Copland. (September 1953)

The Emerging Pattern of International Payments. By Raymond F. Mikesell. (April 1954)

Agricultural Price Policy and International Trade. By D. Gale Johnson. (June 1954)

"The Colonial Sterling Balances." By Ida Greaves. (September 1954)

America's Foreign Trade Policy and the GATT. By Raymond Vernon. (October 1954)

The Bank for International Settlements, 1930-1955. By Roger Auboin. (May 1955)

United States Merchant Marine Policies: Some International Economic Implications. By Wytze Gorter. (June 1955)

International Cost-Sharing Arrangements. By Thomas C. Schelling. (September 1955)

SPECIAL PAPERS IN INTERNATIONAL ECONOMICS

1. A Survey of International Trade Theory. By Gottfried Haberler. (September 1955)