

A. Carreras Marín

*Geographical effects on the accuracy of textile trade data*

Geographical effects on the accuracy of textile trade data:

an international approach for 1913\*

Anna Carreras Marín

Universitat Pompeu Fabra

Universitat de Barcelona

## Summary

Foreign trade statistics are the main data source to the study of international trade. However its accuracy has been under suspicion since Morgernstern published his famous work in 1963. Federico and Tena (1991) have resumed the question arguing that they can be useful in an adequate level of aggregation. But the geographical assignment problem remains unsolved. This article focuses on the spatial variable through the analysis of the reliability of textile international data for 1913. A geographical bias arises between export and import series, but because of its quantitative importance it can be negligible in an international scale.

JEL Classification: B4 N01 N73

Keywords: economic geography, statistical accuracy, economic textile history

\* This paper is financed by two research projects: *Importaciones y modernización económica en América Latina, 1890-1960*. BEC2003-00190MICYT, and

*Industrialización y estrategias empresariales en España, 1780-1936, BEC2002-01113MCYT*

### I- Foreign Trade Statistics in 1913: first efforts in homogenisation versus a panorama of divergence

Foreign Trade Statistics have been broadly used for many studies of international trade before the First World War. Anyway their value as a source of accurate information has not always been accepted even for the period after the Great War. The main problem in their accuracy consists in the lack of homogeneity, which makes impossible any attempt of comparability. The fact is that we have two data series that are supposed to be the same thing. Exports of a country are imports of its trade partner, so these two figures have to match perfectly. But as long as they have been collected from two different countries, they almost never coincide exactly.

Such disagreements are bigger, the more in the past you go. The efforts of international institutions for the generation of a common statistical framework have solved gradually the problem. As it is well known the evolution of international institutions is closely linked to the improvement of international statistics. But what kind of scenario did we have before the First World War?

Data from 1913 are near the first efforts in these attempts at homogeneity. In 1910 it was celebrated in Brussels the *Conférence Internationale de Statistique Douanière*, in which the bases for a first set of common rules for the compilation of Foreign Trade Statistics has been stated. It was the beginning of what it would culminate in the first Standard International Trade Classification (SITC), a basic tool for international trade studies nowadays. In 1913 the intention of homogeneity was starting, however there were serious discrepancies between countries' statistics<sup>a</sup>.

Deep analysis of the nature of these discrepancies had been developed after the Second World War. In 1953, Allen and Elly had described the main causes for such statistical discrepancies. These authors found five reasons for the fact that a country's exports don't coincide with its trade partner's imports, although it must be the same.

The first cause of statistical discrepancies was the different definitions of each commodity. In the textile sector, for example, carpets were aggregated with the finished wool commodities but they also were registered separately in a clothing item<sup>b</sup>. To avoid this problem the focus of this work has been carried out on the textile sector in a very aggregate level, trying to neutralize differences coming from definitions.

A second cause of statistical deviations is the geographical assignment. It comes from the fact that some countries assigned imports to the last harbour where the ship had stopped and not to the country where the commodity had been produced. Otherwise some countries also registered the first destination where their exports had gone and not the final country where the commodity would be consumed. Tena and Federico assumed that the second option was more frequent than the first one, being the exports tendency more geographically concentrated than the imports one as a result. This proximity effect will be tested later in this article.

A third element of statistical discrepancies identified by Allen and Elly is closely related to the first one. It consists in the different levels of aggregation each country used for the compilation of its statistics. In the textile example used above, if the statistic of a

country is not enough detailed it will be impossible to know if carpets are included in the woollen elaborated commodities or in the textile clothing.

All these three statistical discrepancies can appear either on weights or in values. But the fourth cause of differences is only related to the valuing system of each country. Transport costs were usually thought to be included in imports but not all countries followed this convention<sup>c</sup>. Systems of pricing the commodities were also different. Some countries used official values, the more accurate the more recent they were stated. Other countries used declared values, the less tariffs they impose, the more accurate they were. Finally it also interferes the rate of exchange each country used to translate foreign values into its own currency. As long as the textile data of this work was in values and it was impossible to transform it in weights, all these facts have been taken into account.

A last cause of discrepancies refers to the capability of the trade agents to bring the information in an accurate manner. In this sense, discrepancies can be attributed to an intentionally concealment or simply to omissions. Following this theory, duties make more confident the country efforts to obtain information about its imports. But it also arises a big suspicion about the reliability of the information given by an agent, which resulted in him paying for it. As a result of these opposite interests, tariffs have an important ambiguous effect on the accuracy field.

## II- On the accuracy of Foreign Trade Statistics: from pessimism to pragmatically optimism

Allen and Elly had offered a systematic analysis of the causes that can explain discrepancies between trade partners' statistics. Otherwise Morgernstern (1963) elaborated an index to measure the importance of such bilateral discrepancies. This author weighed the absolute difference between the two countries' statistics for the single amount of trade for one of them. This measure is a percentage of one of the two countries' volume trade:

$$(1) \frac{I_1 - E_2}{I_1} \qquad (2) \frac{E_1 - I_2}{E_1}$$

where:  $I_1$ = Imports from country A following A's statistics  
 $E_1$ = Exports from country A following A's statistics  
 $I_2$ = Imports from country B following B's statistics  
 $E_2$ = Exports from country B following B's statistics

Morgernstern's index was applied to different years: 1909/13, 1928, 1935, 1938, 1948, 1952, 1956 and 1960. The countries used for those comparisons were supposed to have the best statistics: United States, Canada, Belgium, Great Britain, Germany and France. Although this sample had to be the best in statistical accuracy, Morgernstern's measures had presented quite pessimistic results. Differences higher than 25% were not unusual between pairs of countries, and there were also a significant group of countries with differences above 50%. On the other hand, discrepancies lower than the 10% were the least, although they have become more important in recent periods of time. Good news for studies focused after the Second World War, bad news for the periods before.

According to Morgernstern, a divergence under 25% could be considered as a good result, behaving really optimistic. The reason was that such discrepancy can be explained by the presence of transport costs or tariffs, assuming that imports were c.i.f. and exports were always f.o.b. If both countries were distant, you could think there were high transport costs included in imports. If those countries bordered each other, you could think that there were high tariffs between them. In both cases, imports should be bigger than exports. The result of this difference would be a negative sign. And then the second big problem arises: the signs weren't in the way they were supposed to.

The pessimistic view had been supported by these two elements: huge differences and arbitrary signs. If no further research was made in this field, we had been forced to forget Foreign Trade Statistics, before the Standard International Trade Classification (SITC) had come to scene. But obviously, the pessimistic point of view has been overcome.

Confidence on the official statistics has been recovered in country level. Researchers have been concentrated in their own country statistics trying to compensate national biases. It has been, for example, the case of Spain with extraordinary high tariffs and big problems coming from ancient official values<sup>d</sup>. Big errors of measure have been identified, too. For example, the inclusion of re-exportations is well known in the case of The Netherlands or Austria-Hungary<sup>e</sup>.

Although bilateral comparability shows quite bad results, it doesn't mean that the two data series are incorrect. So, we can assume that errors are only in one of them. It is wide believed that it exists a positive relationship between economic development and

statistical accuracy. Following this belief, the statistics of Great Britain, Germany, France and United States have gained confidence.

Federico and Tena in 1991 revisited Morgernstern using a systematic serious method. Adopting an international approach, they overcame the country level, evaluating the source as a whole, but without forgetting the particular biases of each country's statistics. They offer an alternative index in which bilateral differences are reduced by adding the data from all trade partners:

$$M_i = (\sum M_{ij} / \sum X_{ji}) \times 100 \quad X_i = (\sum X_{ij} / \sum M_{ji}) \times 100$$

This measure avoids the geographical assignment problem, solving Morgernstern's pessimism. Following this method, they have supported the use of Foreign Trade Statistics even for the period before the Great War. The sample used by these authors covered the 90% of the trade for 19 European and 14 non-European countries in 1909-1913, 1928, and 1935. In essence, they found that individual errors compensated each other by aggregation. As a result, discrepancies are not significant in any period of time.

### III A question remaining: the geographical assignment problem

Federico and Tena made two more interesting contributions. First, they tested econometrically the common convention of a positive relationship between economic development and statistical accuracy. Their results show that it is significant in 1913's exporters data, but not after the Great War. As long as manufactured textile exporters in that period were all the countries with high levels of economic development, this means that exporters' data are as reliable as importers' data.



Aggregation reduces discrepancies for each country as a whole, but it hides individual discrepancies between pairs of countries. Such divergences for each two countries are said to be the biggest. Anyway, it is not still the time to refuse the use of these statistics. Federico and Tena made an advice against the geographical assignment problem, but additionally they also made an alternative proposal. They suggested that probably importer's data would be better than the exporter's, from a geographical point of view.

They pointed out that there could be a general tendency in overstating neighbour countries, understating the long distance ones, into exporter's statistics. It is based on the little interest that countries had in the final destination of their exports, in contrast with the strong interest that some countries had in identifying the true origin of their imports, especially when a tariff was associated. They also included the same possibility for the importer's side. A country could also register the last harbour from which an import had arrived, and not the place where the goods were originally produced. According to this hypothesis there could be a proximity effect in both data series, but tariffs could have eliminated it in the import side.

We can also add a scale effect to explain geographic divergences between exporters and importers. In this case we have to take into account that textile exports were highly concentrated in a few number of countries, meanwhile textile imports were more scattered. As a consequence, exporters could omit some destinations because the volume of their trade was quite small. But we can locate this small trade partners looking up the importers' data. If this hypothesis proves to be right, we can make two assessments. We can forget the geographical bias of the exporters' data in an

international level, because we are losing a really little quantitative information. In that level we will assume that the exporters' data are accurate enough. But if we are focused on a regional level, then we have to complement this information with that coming from the importers'.

If we summarize the hypothesis, we have a puzzle to be solved. Optimism about Foreign Trade Statistics is limited to an aggregated level for each country. Either exporters or importers are unreliable for the geographical assignment question because the amount of differences increases when bilateral trade are taken into account. On the other side, importers are supposed to be more accurate in the geographical origin of trade due to their interest in taxing or because they also collect the smallest quantities. But some importers overvaluated their trade by using c.i.f. values.

If we want to solve this puzzle, we have to verify if textile bilateral trade in 1913 was unreliable enough to any geographical purpose. We use Morgernstern's indexes to measure the relative importance of each bilateral difference between exporters' and importers' data. We have eliminated the smallest quantities, having a matrix of 226 registers.

Results following Morgernstern's method are so pessimistic as those of his. Amounts of discrepancies jump to around 1000 % in two extreme cases, although they are mainly under 100%. Discrepancies below the 25% are only in 80 observations, which means a 35% of all of them. Sign results are not much better: differences are either positive or negative.

Things change completely when we take into account the significance of the absolute amount of each discrepancy in terms of total textile trade. In this case, the bigger differences mean only 1%, corresponding to three bilateral exchanges: France with UK, France with Belgium and UK with Australia. Otherwise, 94% of the observations represent less than 0.5%. As a result we can assume that bilateral trade data are reliable for the study of textile international trade. We pretend to improve these results taking into account the geographical bias, which can be an important cause of bilateral discrepancies.

#### IV Textile trade data for 1913: were exporter's data more concentrated than importer's?

A first step to give an answer to the hypothesis about the geographical bias, it has been the concentration measure of both series. Table 1 shows Gini's concentration index of textile-manufactured exported goods in 1913. In the first column, exports are taken from exporter's statistics. In the second one, they have been calculated from trade partners. If exporters had a geographical bias, the first column should be bigger than the second one.

(Table 1 should be here)

In a world aggregate level, we have to assume that exporters' data are more concentrated than importers'. But in the country level, we find a more diverse panorama. For eight individual countries, concentration is exactly the same, meaning that there isn't any geographical bias in exporters or that the bias is in both of them. But

in all of these cases, high values of Gini's index indicate that exchanges included are not more than 1. So, they have not any relevance.

It occurs just the contrary as expected for two countries. This is especially surprising since Switzerland is one of these unexpected cases. Swiss textile exports represented in 1913 a 5% of the world total. It was the fourth exporter after Great Britain, France and Germany. High concentration of the data from Swiss trade partners, it can easily be explained by the fact that Switzerland has not a direct exit to the sea. As long as Swiss textile goods had to be sent through other countries, they could be consigned to these other countries by importers. In this case, there will be a geographical bias but in the importer's side.

As we look down in table 1, the distance between exporter and importer's concentration increases. France, Germany, Great Britain and United States have data more clearly concentrated. This is even more important if we take into account that these four countries represented more than 70% of the world textile trade in 1913. A geographical bias seems to have been identified but we still don't know if it corresponds to a proximity bias or a scale effect.

#### V The spatial pattern of textile export data: is there a scale effect?

Measures of concentration, such as Gini's index, are in fact very weak to catch the spatial dimension. Geographers have developed a method to understand how variables are located in the space: they are called maps. We have mapped textile exports of five countries where exporter's data was more concentrated: USA, UK, Germany, France,

and Italy (see fig. 1-6). We have also included the unexpected case of Switzerland. In this sample we have 78.61 % of total textile trade. We have represented each country as a circle, in order to neutralize visual confusions caused by the different size each country has.

We have drawn three kinds of circles: if a country's exports are only recorded in the exporter statistic, only in the importer one, or in both of them. If there was a scale effect we will find most of the only importers' circles located in many small countries. It has to be remained that we are forgetting for a moment the amount of each difference.

These qualitative maps show three important facts. First, they confirm Gini's results clearly. Importers' data are more dispersed than exporters' data. This means that the use of exporters' statistics implies a missing of geographical information in all cases. Second, these missing circles are not arbitrarily distributed. They are concentrated in Latin America and Africa. As a consequence, any study of these two regions has to take care of using only exporters' data.

Third, many of the missing circles correspond to small countries. We can identify a scale effect, especially for Germany, France and Switzerland. Although importers' data are always more dispersed, the other countries seem to be more geographically exhaustive.

VI The spatial pattern of textile exporters' data: is there a proximity bias?

In order to view the importance of each difference we have mapped its values and signs instead of the previous circles (see fig. 7-12). Now these maps represent the magnitude of the geographical phenomenon as well as they show some clues about its nature. A positive figure may be a redistribution point, as long as exporters' data is bigger than the importers' one. If positive signs are clustered near the exporter country, then we will have a proximity effect in its data. On the other hand, a negative sign could be attributed to transport costs, to high tariffs or even to errors in the importer side.

These quantitative maps show that exporters' data is not only more concentrated but closer to the exporter. A proximity effect is clearly detected in all countries, with no exception. But some overstated data come from a redistribution role, and not from a proximity effect. It is the case of Argentina (except for USA and Germany), Cuba (except for France, Italy and Switzerland) or British India (except for USA and Germany).

This result reinforces the previous one: importers' data was better from a geographical point of view. We can also give an explanation of one of Morgernstern's problems: a positive sign of the difference between exporters' and importers' data is a consequence of a proximity bias in the first country or a redistribution role in the second one. But we don't know the scope of these findings yet.

#### VII The importance of the geographical bias: how big is big?

We have identified two kinds of geographical biases from the exporters' series: one caused by a scale effect and the other one by a proximity or redistribution effect. Its

importance will not be the same depending on the scale of each study. In an international level, such deviations can be omitted, but in the case of some particular regions they are so important that importers' data has to be included.

For the study of textile international trade, the use of exporters' statistics means a major guarantee in terms of their values but it also implies a lack of information about some destinations. Table 2 shows the scope of losing this information for each country and world totals. It has to be considered that these are maximum figures, because import data includes tariffs and transport costs, in most of the cases.

(table 2)

Omitted textile trade in the exporters' statistics represents as a maximum less than 5% of total world volumes. UK' omissions were only 2% in the international level, being the biggest of all countries. In a regional level, only the case of USA appears to have really importance. For this country missing data represented a bit more than 15%. Germany was around 9% and UK was near 5%. All the others were below such percentages.

With the exception of USA, these shares aren't big enough to invalidate the exporters' data. As a result, in terms of geographic diversity we can assume exporters were accurate enough in an international level and for the main countries. These results will not be the same for some of the importer regions, such as Latin America or Asia.

We have stated that higher geographic diversity of importer data is not quantitative relevant, except for the USA. How about the proximity bias? In table 3, we have calculated the cluster of positive differences related to the exporter and to the world. This is a measure of neighbour countries overvaluation. As a result we observe that the importance of the proximity effect is really low in world totals, representing a 5%. At a regional level, it is only significant for USA and France. In the American, Swiss and British cases, the amount of missing trade is almost exactly the same as the amount of overstated exportation. We can assume that these two figures compensate each other. For the other countries, things don't seem to be so clear.

(table 3)

To conclude, we have detected a geographical bias in the exporters' data series coming from three different phenomenons. According to Federico and Tena (1991) we have found a proximity bias. We have also added a scale effect and a redistribution bias. In an international scale none of them are quantitative relevant, but they have different impact in the regional level. In these later cases, exporters' information will be better complemented with that coming from importers. Being optimistic, we can validate Foreign Trade Statistics as a source for geographical studies, provided all biases detected. Exporters' data can be used for geographical purposes, but in some cases they must be complemented with some information coming from importers.



References

- Allen, R. and E. J. Elly, *International Trade Statistics* (New York, 1953).
- Don, Y., 'Comparability of International Trade Statistics: Great Britain and Austria-Hungary before World War I', *Economic History Review*, XXI, pp. 78-92, (1968).
- Federico, G. and A. Tena, 'On the accuracy of foreign trade statistics (1909-1935): Morgernstern revisited', *Explorations in Economic History*, 28, (1991), pp. 259-273.
- Kertesz, A., *Die Textilindustrie sämtlicher Staaten*, (Viewig&Son, Braunschweig, 1917).
- Morgernstern, O., *On the accuracy of economic observations*, (Princeton. New Jersey, 1963).
- Tena, A., 'Las estadísticas históricas del comercio internacional: fiabilidad y comparabilidad' Banco de España. Servicio de Estudios de Historia Económica, 24. (1992)
- Tyszynski, H., 'World trade in manufactured commodities, 1899-1950', *Manchester School*, (sept. 1951), pp. 272-304.
- Société des Nations, *Mémorandum sur le commerce international et sur les balances des paiements, 1912-1926*, (Géneve, 1928).

## Tables

Table 1- Gini indexes for textile exports in 1913

Countries	own exports	others' imports	Countries	own exports	others' imports
Switzerland	0,88	0,91	Japan	0,93	0,92
Netherland India	0,97	1,00	Belgium	0,88	0,86
Bulgaria	0,99	0,99	British India	0,90	0,88
Portugal	0,99	0,99	Persia	0,98	0,96
Russia	0,98	0,98	Spain	0,92	0,89
South Africa	1,00	1,00	Turkey	0,96	0,92
French Asia	0,98	0,98	Italy	0,81	0,77
China	0,89	0,89	France	0,89	0,84
Australia	1,00	1,00	Germany	0,83	0,75
New Zealand	1,00	1,00	Great Britain	0,77	0,69
Austria-Hungary	0,88	0,87	USA	0,92	0,84
The Netherlands	0,96	0,95	<b>World</b>	<b>0,87</b>	<b>0,84</b>

Source: Gini indexes elaborated from Kertesz original data (1917)

Table 2- The scope of losing some geographical information

	% Exporter	% World
USA	15,69	0,33
GER	9,14	1,35
UK	5,21	2,05
FRAN	4,23	0,64
ITA	2,88	0,12
SUI	3,05	0,15

Source: Elaborated from Kertesz data (1917)

Table 3- The scope of overstating neighbour exports

	% exporter	% world
USA	15,67	0,34
GER	5,77	0,86
UK	5,32	2,09
FRAN	9,62	1,45
ITA	6,04	0,26
SWI	3,89	0,19

Source: Elaborated from Kertesz data (1917)

Qualitative maps

Fig. 1- Textile exports of USA

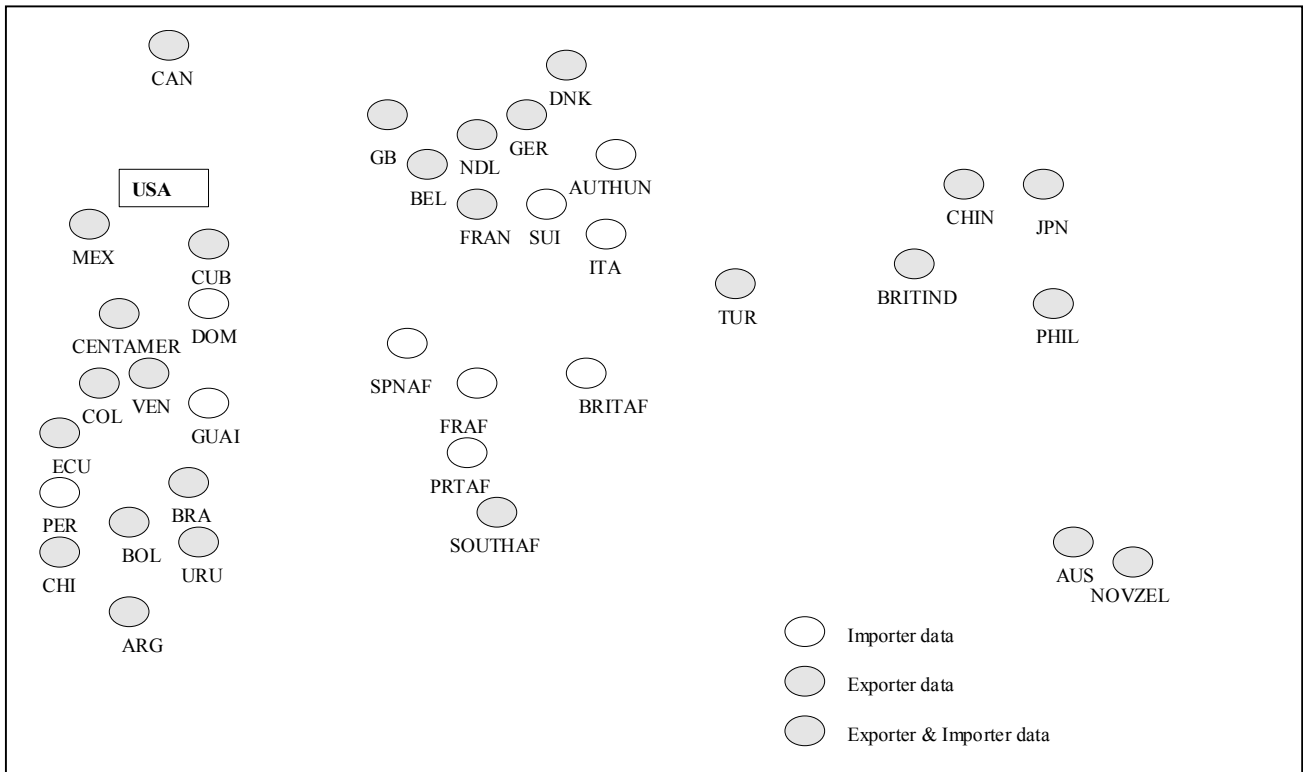


Fig. 2- Textile exports of Great Britain

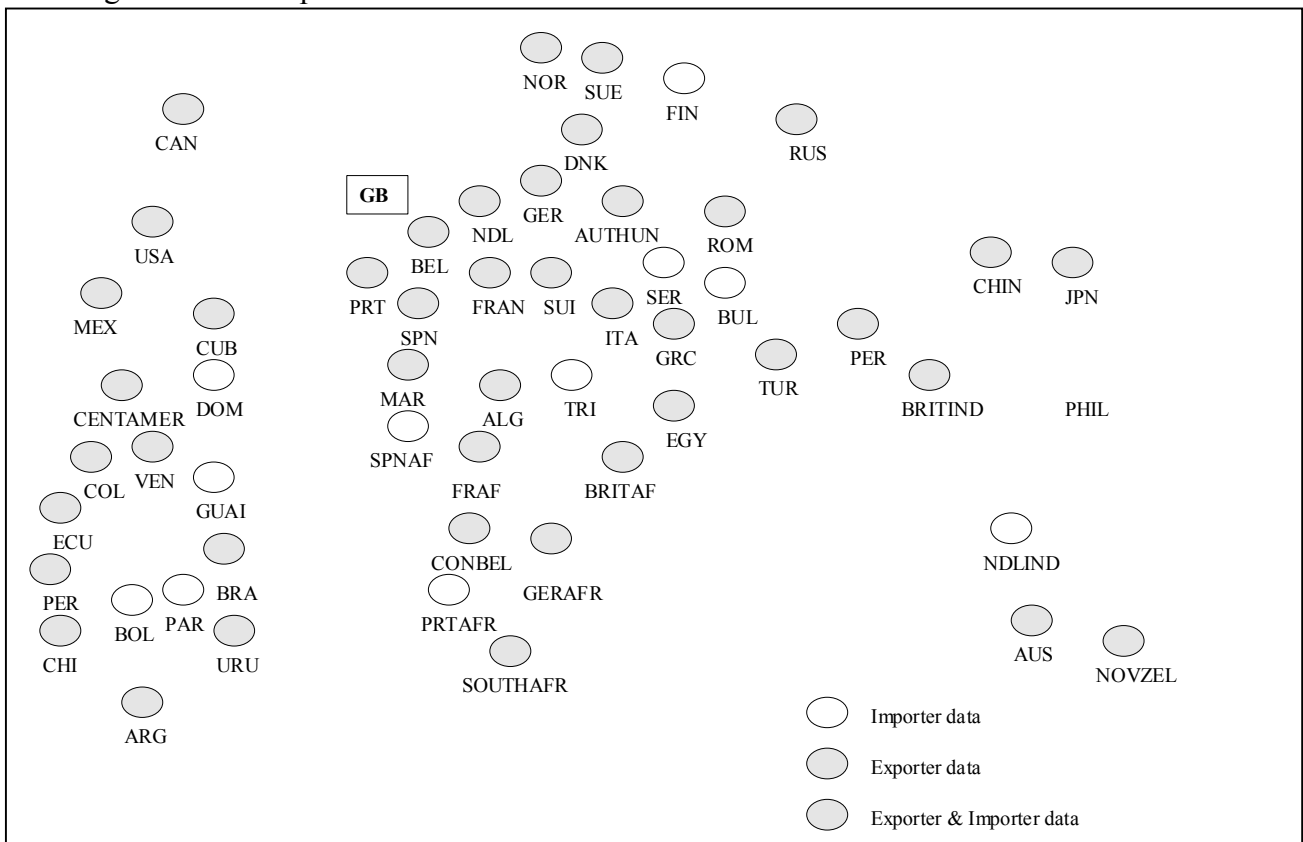




Fig. 5- Textile exports of Italy

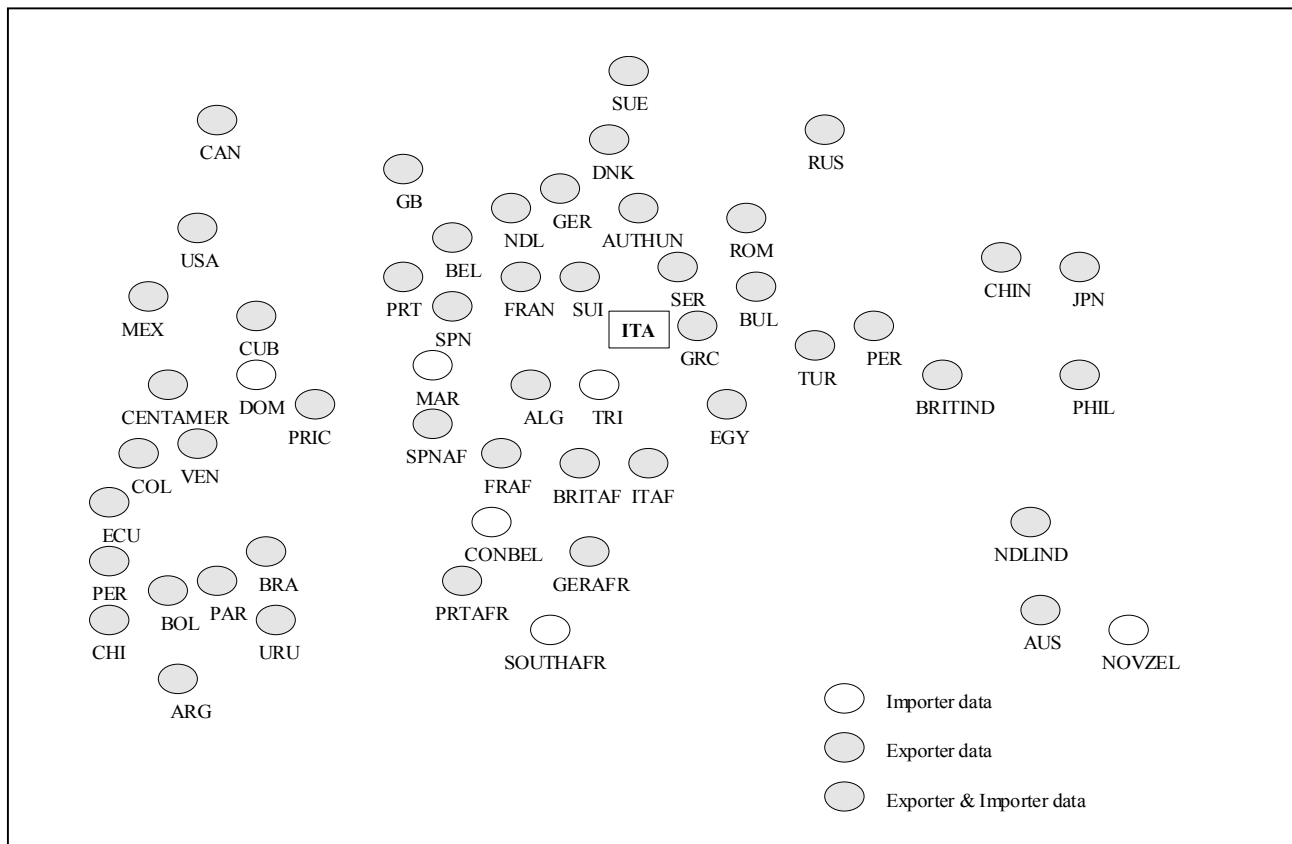
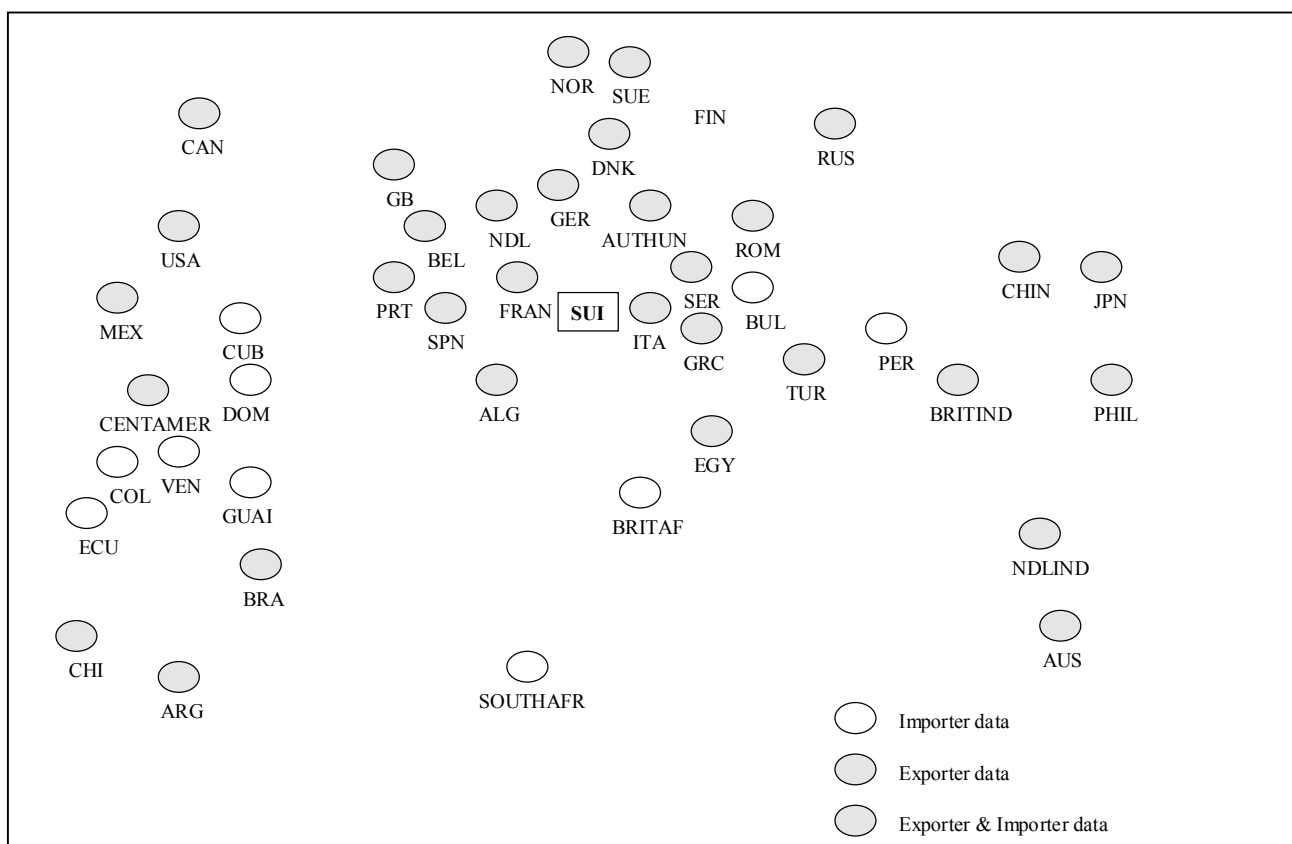


Fig. 6- Textile exports of Switzerland



**Quantitative maps**

Fig. 7- Statistical differences of USA's textile exportation

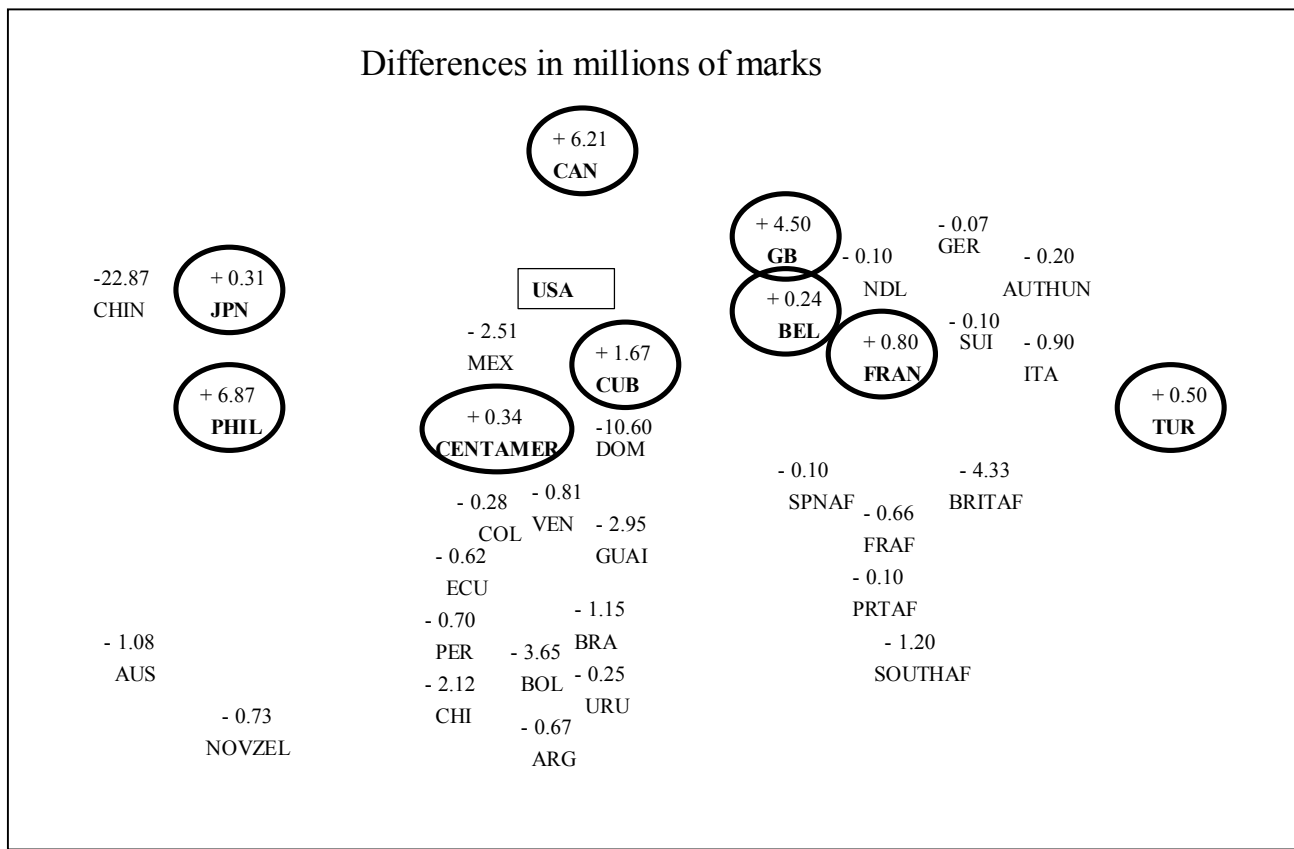


Fig. 8- Statistical differences of GB's textile exportation

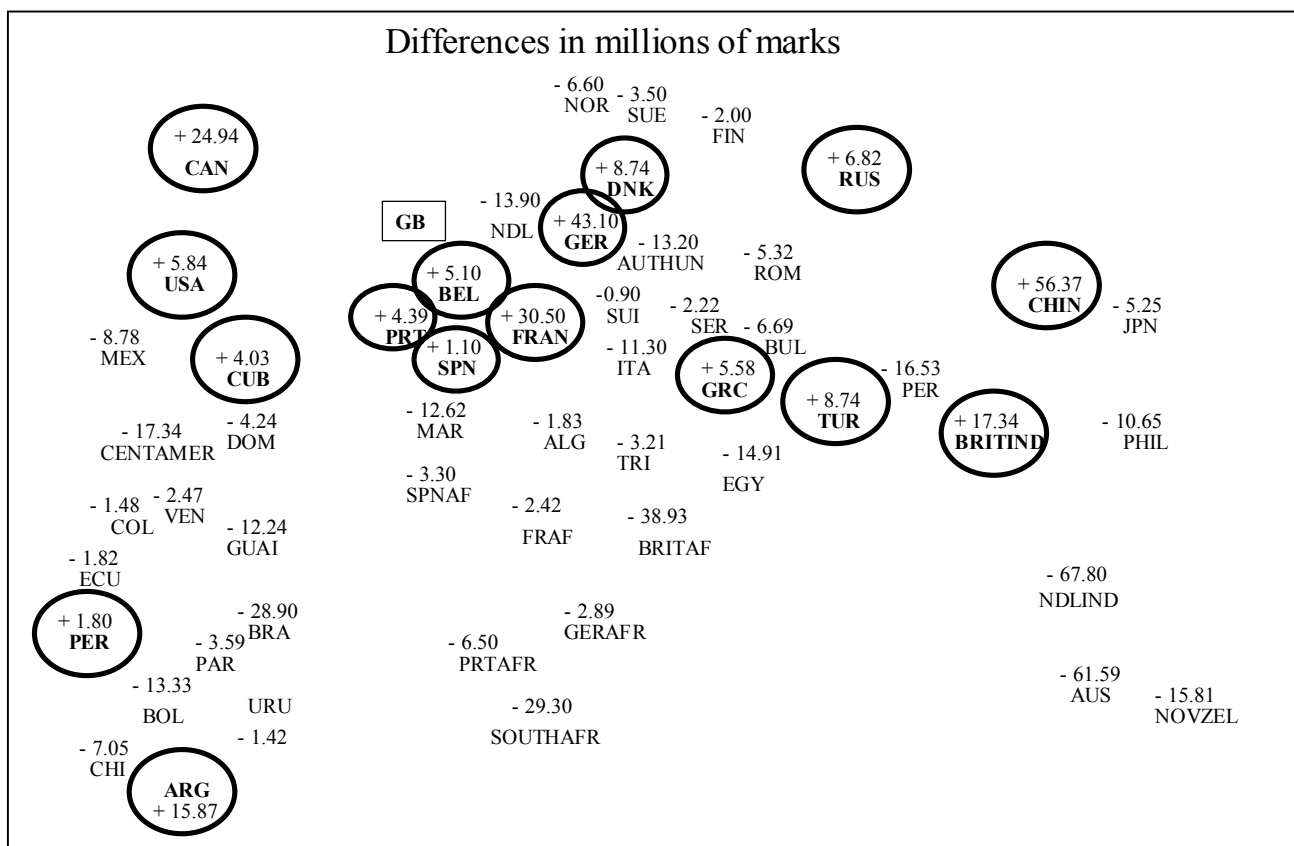


Fig. 9- Statistical differences of Germany's textile exportation

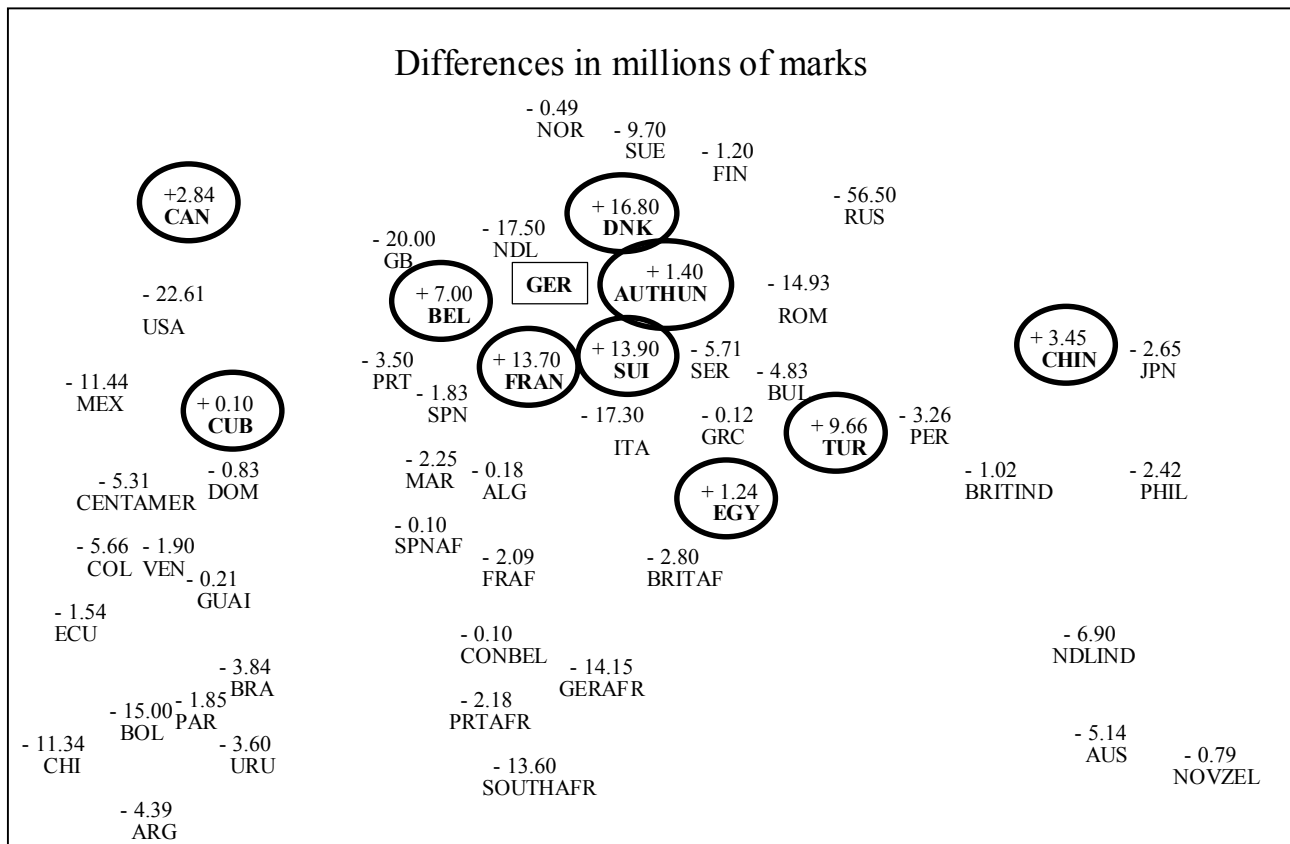


Fig. 10- Statistical differences of France's textile exportation

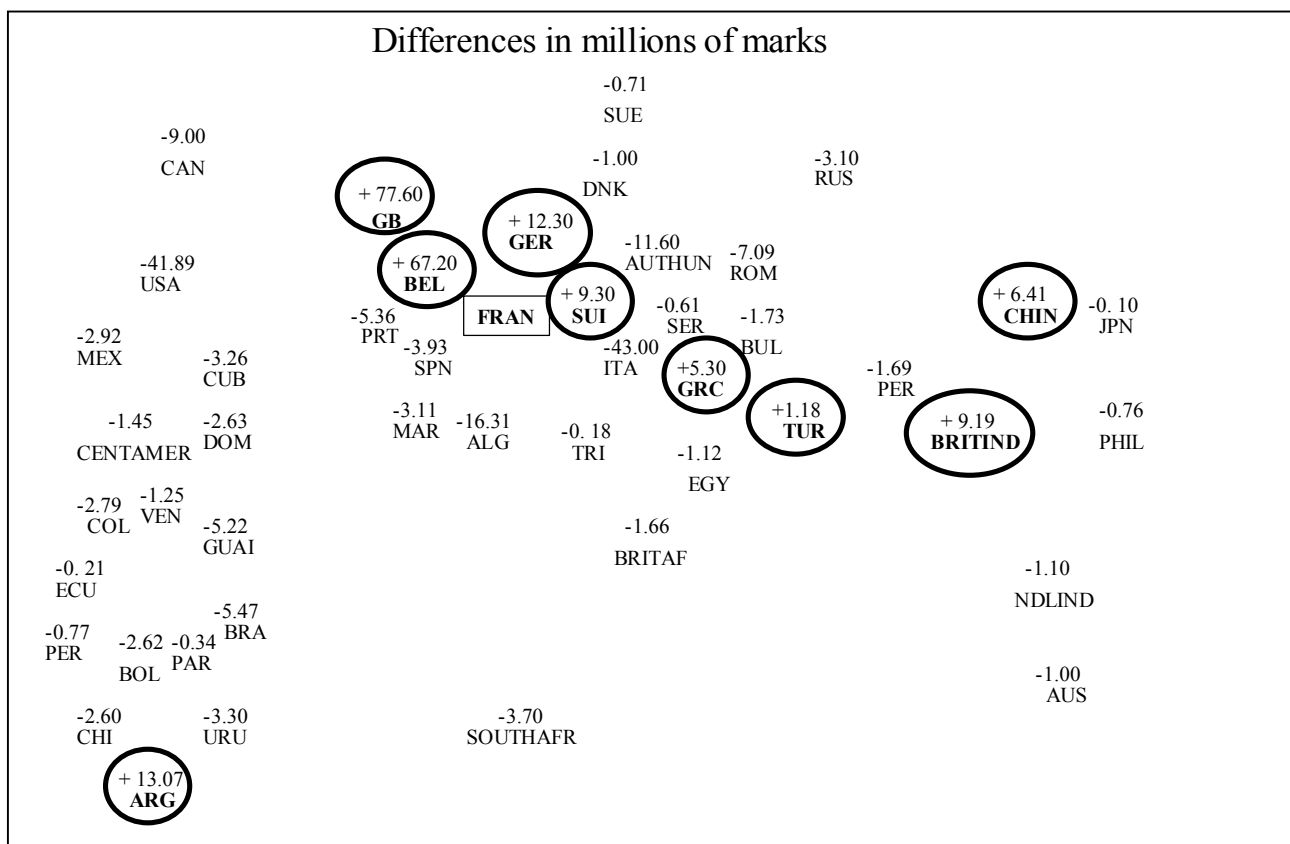


Fig. 11- Statistical differences of Italy's textile exportation

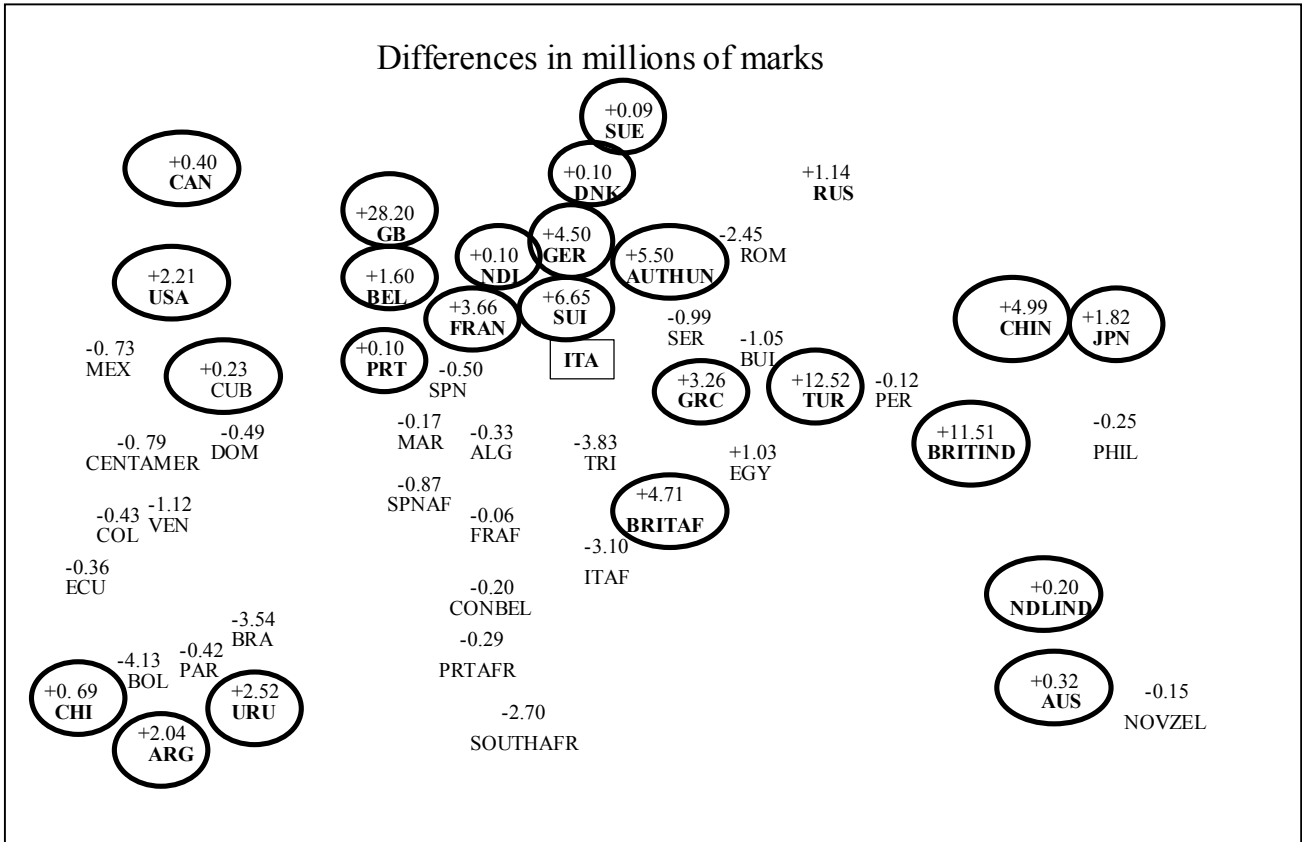
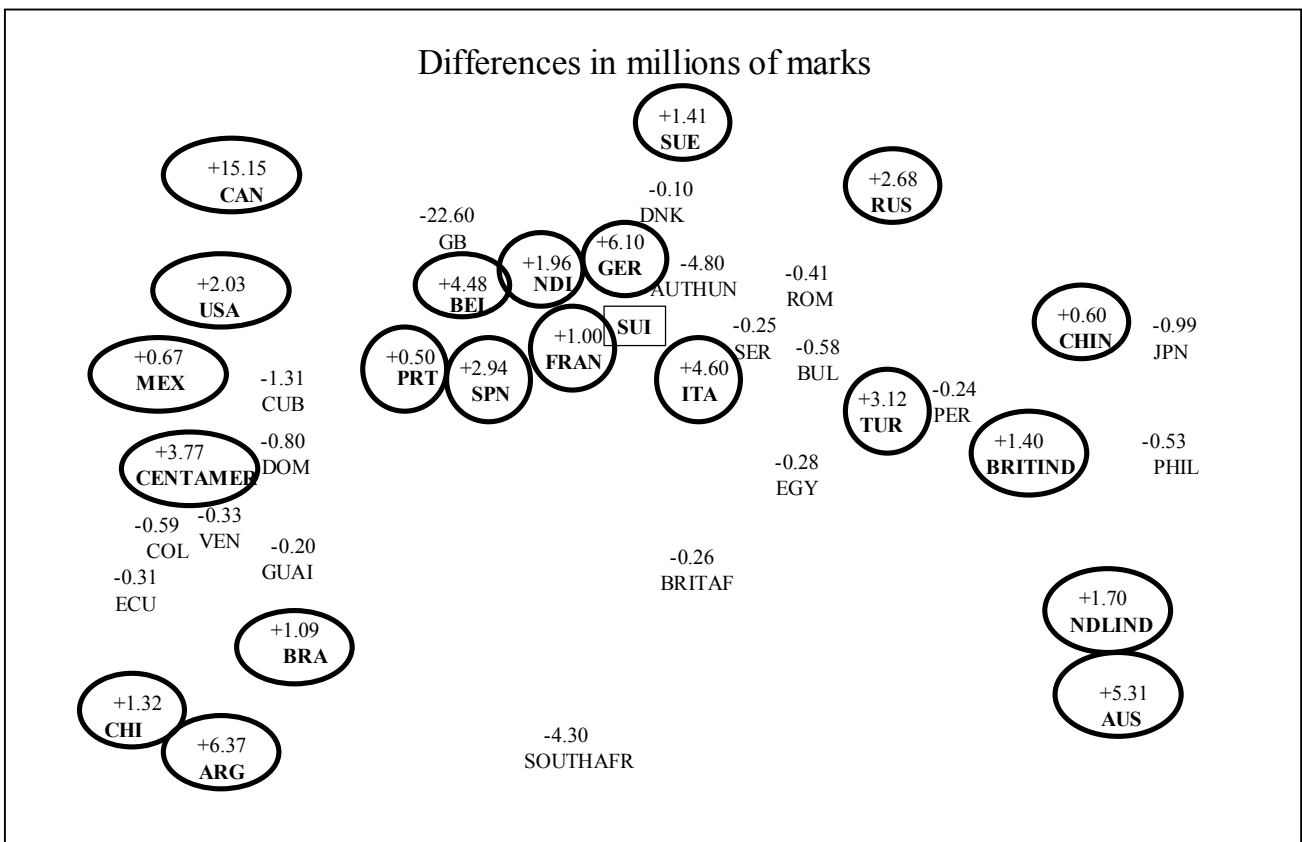


Fig. 12- Statistical differences of Switzerland's textile exportation





<sup>a</sup> Société des Nations (1928)

<sup>b</sup> In the clothing item, all kind of textile fibres was included.

<sup>c</sup> In 1913, Australia, Canada, Chile, Cuba, Rep. Dominican, USA, Guatemala, Honduras, Mexico, Nicaragua, Panama, Russia, Salvador, South Africa, and Venezuela used f.o.b. imports. (Société des Nations, 1928, pp. 26-27)

<sup>d</sup> Tena (1991)

<sup>e</sup> Tyszynski (1951), Don (1968).