Comparing the Networks of Ethnic Japanese and Ethnic Chinese in International Trade

著者	Kumagai Satoru
権利	Copyrights 日本貿易振興機構（ジェトロ）アジア経済研究所 / Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) http://www.ide.go.jp
journal or publication title	IDE Discussion Paper
volume	113
year	2007-07-01
URL	http://hdl.handle.net/2344/642

INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 113

Comparing the Networks of Ethnic Japanese and Ethnic Chinese in International Trade

Satoru KUMAGAI*

Abstract

In this paper I re-examined the trade enhancing effects of ethnic Chinese networks, found by Rauch and Trindade (2002), on a newer and extended data set. The effects are estimated by the gravity equation with the product of the population ratio (or absolute number) of the ethnic Chinese in both the importing and exporting countries, and are reaffirmed positive and statistically significant. I also compared the effects of two different ethnic Japanese networks, i.e., the networks of long-term Japanese stayers in foreign countries, and the networks of permanent Japanese residents in foreign countries. It is found that the former has stronger trade enhancing effects than the latter. This shows that the effects of ethnic networks on international trade can be generalized beyond the ethnic Chinese, and the 'cohesiveness' of the ethnic network matters to the trade enhancing effects of the network.

Keywords: trade, networks

JEL classification: F10

---

*Researcher, Economic Integration Studies Group, Inter-disciplinary Studies Center, IDE (kumagai@ide.go.jp)

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

Comparing the Networks of Ethnic Japanese and Ethnic Chinese in International Trade

Satoru KUMAGAI*

Abstract

In this paper I re-examined the trade enhancing effects of ethnic Chinese networks, found by Rauch and Trindade (2002), on a newer and extended data set. The effects are estimated by the gravity equation with the product of the population ratio (or absolute number) of the ethnic Chinese in both the importing and exporting countries, and are reaffirmed positive and statistically significant. I also compared the effects of two different ethnic Japanese networks, i.e., the networks of long-term Japanese stayers in foreign countries, and the networks of permanent Japanese residents in foreign countries. It is found that the former has stronger trade enhancing effects than the latter. This shows that the effects of ethnic networks on international trade can be generalized beyond the ethnic Chinese, and the 'cohesiveness' of the ethnic network matters to the trade enhancing effects of the network.

* Researcher, Economic Integration Studies Group, Inter-disciplinary Studies Center, IDE (kumagai@ide.go.jp)

1. Introduction

International trade is anything but 'frictionless.' The recent progress in international trade theory has revealed that trade costs have significant effects on the volume and the pattern of international trade. For instance, the gravity equation, which has a notably successful track record in empirical literature, shows that the volume of bilateral trade decreases as the distance between the trading partners, i.e., the trade costs between the two countries, increases.

Although transport costs make up a significant portion of international trade costs, other kinds of trade costs are also important and have increasingly drawn attention. Rauch (1996, 1999) asserts that search costs prevent traders from extracting trade opportunities to their full potential. Rauch proposes two ways of overcoming the 'uninformativeness' of goods and trade partners. The first way is to use organized commodity exchanges, if the goods are homogeneous. The other way is to use ethnic, family or corporative networks, mainly for trading differentiated goods. Rauch and Trindade (2002) show that ethnic Chinese networks have economically important impacts on the volume of international trade.

In this paper I compare ethnic Japanese networks and ethnic Chinese networks in international trade following the method and procedure used by Rauch and Trindade (2002). More precisely, I have estimated the gravity equation with the product of the ratio (or absolute number) of the Japanese (Chinese) population in the trading partners, and tested the statistical significance and magnitude of the coefficients. This study will be a valuable contribution to the yet scarce empirical research based on the ‘search/matching’ view of international trade presented by Rauch (1996, 1999).

This study has three main advantages over Rauch and Trindade (2002) and other preceding research. First, I have estimated the effects of both ethnic Japanese and ethnic Chinese networks on international trade together. This enables testing whether the effects of ethnic networks on international trade can be generalized beyond the ethnic Chinese. Second, I have estimated the effects of two different kinds of Japanese networks: the networks of long-term Japanese stayers in foreign countries and the networks of permanent Japanese residents in foreign countries. This comparison makes it possible to tell how the ‘cohesiveness’ of networks matters for the effects on international trade. Third, I utilize a newer and more extensive data set (139 countries in 2000) compared with Rauch and Trindade (63 countries in 1980 and 1990). The comparison between the coefficients on the data set as of 2000 and those as of 1990 enables us to analyze how the progress of 'globalization' in the 1990s has affected the importance of ethnic networks in international trade. In addition, the more than doubled number of countries improves the reliability of the estimates.

This paper is organized as follows. Section 2 briefly reviews the previous research on the search/matching view of international trade, which is background to the estimations in this paper. In section 3 the methods used in this paper and the hypotheses to be examined are presented along with the descriptive information on the two types of ethnic Japanese networks. Section 4 explains the specifications of the gravity equation and data used in the subsequent estimations. The results of the estimations are presented in section 5. The estimations are conducted using several different specifications to test the robustness of the results. Section 6 concludes this paper with a discussion of the findings and some remaining problems.

2. Related literature

2.1 Theoretical literature on the search/matching view

In the real world, it is clear that 'uninformativeness' amongst trading partners obstructs the efficient trading of goods. This is true for domestic trade, but much more so for international trade. The information needed to trade goods internationally is quite extensive. It ranges from the price, quantity, quality and delivery date of goods, the mode and costs of transportation, to the credit worthiness of trading partners.

Despite the importance of information in international trade, with the exception of Rauch (1996,1999), there has been little research that incorporates information issues into the theory of international economics. Rauch presented the 'network/search view' of trade, to address these information issues. This network/search view of trade seems to provide clues to solve some 'puzzles' in international economics.

Rauch (1996) shows that the puzzling relationship between the volume of trade and the distance between trading partners found by Frankel et al. (1993) can be solved by the network/search view of trade. Frankel et al. found that the volume of trade in goods with higher transport costs does *not* decrease more rapidly than in goods with lower transport costs, as the distance between trading partners increases. Rauch explains that this seemingly strange relationship of the volume of trade, trade costs and the distance between trading partners can be explained by the network/search view.

> [D]ue to their heterogeneity, low transport cost commodities are traded through networks while high transport cost commodities either have organised exchanges or reference prices that facilitate international commodity arbitrage.(Rauch 1996, p.2)

Goods with higher transport costs, which generally coincide with homogeneous goods, tend to be traded on organized exchanges. Their volume of trade is less affected by the distance between trading partners because of the efficiency of organized exchanges. Goods with lower transport costs generally coincide with differentiated goods and tend to be traded through various networks. Their volume of trade is likely to be affected by the distance between trading partners because the efficiency of networks declines with the distance. Thus, the trade costs (including search/matching costs) of goods with higher transport costs do not necessarily increase more rapidly than those of goods with lower transport costs, as the distance between trade partners increases.

Apart from the puzzle proposed by Frankel et al., the network/search view of trade seems to provide important clues for solving another puzzle in international trade. For instance, Harrigan (2001) points out that there are a lot of ‘zeros’ observed in disaggregated bilateral trade data. This stylized fact can be explained by the search/matching view of trade. It can be explained that the uninformativeness of goods and trading partners beyond national borders prevents some goods from being traded internationally.

Recently the network/search view has gained more solid theoretical background, and is becoming a full-fledged economic theory. Rauch and Casella (2003) have given the network/search view a mathematical formalization. In the model, international matching proceeds on a circular space where producers are equally distributed at even intervals, and the surplus from the trade is divided equally between trading partners in accordance with the Nash bargaining solution.

2.2 Empirical literature on the network/search view

Although the network/search view of trade provides important clues for solving the puzzles in international trade, little empirical research has been done based on this view of trade. Rauch (1996) refers to Gould (1994) and Brainard (1993), relating them to the network/search view.

Gould (1994) showed that the bilateral volume of trade between the US and other countries is positively correlated with the number of immigrants to the US from their trading partners. This is an example of ethnic networks facilitating bilateral trade by providing more information. However, Gould's immigrant effect is a different concept from the ethnic networks examined in this paper, because Gould limited the networks to only between the home and host country of the immigrants.

Brainard (1993) showed that intra-firm trade is roughly equal to one-quarter of both imports and exports in bilateral trade involving the US. This can be interpreted as corporate networks providing the information needed for the trading of goods.

Rauch (1999) himself showed that the networks formed by common languages and/or colonial ties are more important for the trade of differentiated goods than for homogeneous goods traded on organized exchanges. He also showed that the trade barriers caused by search costs are higher for differentiated goods than homogeneous goods.

Rauch and Trindade (2002) found that ethnic Chinese networks have economically important impacts on the volume of international trade. Their 2002 article is the starting point of this paper and is examined thoroughly in the following sections.

Other than Rauch, Feenstra et al. (2002) showed that Hong Kong provided international buyers with substantial information on export goods from China in the 1990s. They showed that Hong Kong intermediated about half of the goods exported from China to the rest of the world, and the value of intermediary services provided by Hong Kong equaled about 16% of the China's total export value through Hong Kong. This study provides evidence that reliable information is indispensable for the trade in goods with less developed countries.

3. Approach and Hypotheses

3.1 Approach of this paper

Rauch and Trindade (2002) show that the importance of ethnic Chinese networks is higher in the trading of differentiated goods compared with other goods. This result is in line with the theoretical prediction. Because differentiated goods have higher uninformativeness, trading in these goods needs networks to overcome these trading barriers.

I have further tested the viability of the search/matching view of international trade by estimating the effect of ethnic Japanese networks on international trade as well as the effect of ethnic Chinese networks. More precisely, I try to estimate the gravity equation with the product of the ratios (or absolute number) of the Japanese (Chinese) population in the trading partners and test the statistical significance and magnitude of the coefficients.

In estimating these network effects, I have followed the methods and procedures in Rauch and Trindade (2002) as much as possible in order to make the results comparable to those in their original paper. I use the same 'threshold' Tobit estimator, and conduct the estimations using their same specifications whenever possible.

However, I have modified their methods somewhat which leads to three main advantages over their original paper. First, I have estimated the effects of both ethnic Japanese and ethnic Chinese networks on international trade simultaneously. This makes it possible to test whether these ethnic network effects on international trade can be generalized beyond the networks of ethnic Chinese.

Second, I have estimated the effect of two different Japanese networks, i.e., the networks of long-term Japanese stayers in foreign countries (which I call '**modern**' Japanese networks) and the networks of permanent Japanese residents in foreign countries (which I call '**traditional**' Japanese networks). The networks of long-term Japanese stayers are supposed to have higher cohesiveness than the networks of permanent Japanese residents in foreign countries because the former are being continuously regenerated while the latter were formed mainly before World War II. Thus, this comparison makes it possible to test whether the cohesiveness of networks matters for their effects on international trade.

Third, I have utilized a newer and more extensive data set (139 countries in 2000) compared with Rauch and Trindade (62 countries in 1980 and 1990). The comparison between the coefficients on the data set in 2000 and those on the data set in 1990 enables an analysis of how the progress of 'globalization' in the 1990s has affected the importance of ethnic networks in international trade. In addition, the more than doubled number of countries improves the reliability of the estimates.

The methods and data set used in this paper are explained in detail in the next section.

3.2 Hypotheses to be examined

The (null) hypotheses to be examined in this paper are as follows:

Hypothesis 1a The networks of ethnic Chinese *do not* have statistically significant effects on the volume of bilateral trade.

Hypothesis 1b The networks of ethnic Chinese *do not* increase the trade of differentiated goods more than that of homogeneous goods.

Hypothesis 2a The 'modern' networks of ethnic Japanese *do not* have statistically significant effects on the volume of bilateral trade.

Hypothesis 2b The 'modern' networks of ethnic Japanese *do not* increase the trade of differentiated goods more than that of homogeneous goods.

Hypothesis 3a The 'traditional' networks of ethnic Japanese *do not* have statistically significant effects on the volume of bilateral trade.

Hypothesis 3b The 'traditional' networks of ethnic Japanese *do not* increase the trade of differentiated goods more than that of homogeneous goods.

Hypothesis 1 is exactly the same hypothesis discussed in Rauch and Trindade (2002) and which they successfully rejected. I re-examine the same hypothesis here on the newer and more extensive data set. If Hypothesis 1a is not rejected, it can be interpreted that the importance of ethnic Chinese networks vanished during the 1990s. If both Hypothesis 1a and 1b are rejected, then the network/search view of international trade gains another supporting example.

Hypothesis 2 is the main hypothesis discussed in this paper. If Hypothesis 2a is not rejected upon rejection of Hypothesis 1a, then it can be concluded that the importance of ethnic Chinese networks on international trade is exceptional among the many other ethnic networks. Such a result will be somewhat 'damaging' to the network/search view because Rauch repeatedly takes the networks of Japanese *Sogo-shosha* as an example of the networks that provide searching activities in international trade[1]. On the other hand, if both Hypothesis 1a and 2a are successfully rejected, then it can be concluded that the network/search view of international trade can be generalized to other ethnic networks besides those of ethnic Chinese.

Hypothesis 3 is tested here in order to ensure that the rejection of Hypothesis 2 concerns specifically the importance of the 'modern' networks of ethnic Japanese in international trade. As will be mentioned in the following subsection, 'traditional' networks of ethnic Japanese have lower cohesiveness than the 'modern' networks, and are supposed to be unimportant in international trade. Thus, if Hypothesis 3 is also rejected upon rejection of Hypothesis 2, then it can be suspected that the rejection is the result of some other reasons. For instance, ethnic Japanese may live in countries that have similarities in tastes with Japan, and those similarities in tastes increase the trade among those countries. If Hypothesis 3 is not rejected upon rejection of Hypothesis 2, then this results fits the prediction of the search/matching view of international trade, and the cohesiveness of the ethnic networks matters in international trade.

3.3 'Traditional' Japanese networks

[1] Although the networks of Sogo-shosha are not exactly ethnic networks, they overlap generally with the networks of ethnic Japanese. Until recently, many Japanese multinational companies were reluctant to employ non-Japanese in executive positions mainly because of the difficulties in communication with the headquarters in Tokyo where Japanese is the primary language.

To understand the effects of ethnic Japanese networks on international trade, it is helpful to distinguish two kinds of networks, i.e., the 'traditional' networks and the 'modern' networks of ethnic Japanese. In this paper the traditional networks are defined as the networks of overseas Japanese who have permanently left their home country. On the other hand, the modern networks of ethnic Japanese are defined as those of overseas Japanese who have departed from Japan temporarily, mainly because of business requirements.

The overseas Japanese who comprise the traditional networks are mainly the second and later generations of Japanese who emigrated from Japan to foreign countries before the Second World War. They are comparable to the majority of Chinese immigrants because both groups permanently left their home country to seek better economic opportunities. However, traditional networks of ethnic Japanese are not often mentioned in economic literature because they have kept a low profile in business activities. This is partly because the number of Japanese immigrants (less than 300 thousands) is much smaller than that of Chinese immigrants (about 30 millions).

Another reason for the inactiveness of ethnic Japanese traditional networks is their low cohesiveness. Compared with Chinese immigrants, Japanese immigrants tend to assimilate into the host countries. This tendency is partly because of the cultural background of ethnic Japanese[2], and partly because of the sparseness of their population in host countries.

[2] a famous Japanese proverb says that *Go ni ireba go ni shitagae* (you should behave as the others do when you are in a foreign country).

3.4 'Modern' Japanese Networks

The 'modern' networks of ethnic Japanese have completely different characteristics from the traditional networks. The overseas Japanese who comprise the modern networks are mainly the employees of Japanese multinational corporations. The data from the Ministry of Foreign Affairs, Japan (2000) shows that 58.0% of these overseas Japanese are employees of multinational corporations, whereas 24.8% are students and researchers, and 4.8% are government officers and their families. The length of staying in host countries varies from several months to over 20 years. By and large these employees have no obligation to settle in the host countries after their retirement, so their tendency to assimilate into the host country is relatively low compared with permanent Japanese residents.

Anecdotally, most of the children of *Chu-zai in* (overseas Japanese employees) go to *Nihon-jin Gakko* (Japanese School) for fear of failing the entrance examination to university in Japan. *Chuzai-in Zuma*s (wives of overseas Japanese employees) form a community which exclusively consists of *Chizai-in Zuma*s having limited contact with the host country. Thus *Chuzai-in*s and their families are extremely well informed about what is going on in Japan.

Thus the importance of modern ethnic Japanese networks in international trade is supposed to be higher than that for traditional ethnic Japanese networks. The former largely overlap the networks of Japanese multinational corporations, and their high cohesiveness in the host countries is supposed to enhance formal and informal exchanges of trade enhancing information.

4. Model and Data

4.1 Theories behind the gravity equation

While the gravity equation has a very good track record within a large body of empirical research, there is frequent criticism that the theory behind the equation is not clear. However, the problem is not that the gravity equation has no theory behind it, but that it can be derived from many different theories (for instance, Anderson [1979], Harrigan [2001], Feenstra et al. [2001]), and it cannot be determined exactly which theory is behind the empirical success of the gravity equation.

Rauch and Trindade (2002) avoid the derivation of the gravity equation from any theoretical model by quoting Deardorff(1998).

> [A]ny plausible model of trade would yield something very like the gravity equation, whose empirical success is therefore not evidence of anything, but just a fact of life (p. 12)

However, I will briefly show how the gravity equation is derived from the combination of the Armington assumption, monopolistic competition and the CES utility function, following Anderson and Wincoop (2002).

Country j's consumption of the goods from country i is determined by solving the following maximization problem.

$$\arg\max_{c_{ij}} = (\sum_i \beta_i^{1/\sigma} c_{ij}^{(\sigma-1)/\sigma})^{\sigma/(\sigma-1)} \ \ s.t. \sum_i (p_i t_{ij}) c_{ij} = y_j \ \ \ldots(1)$$

where σ is the elasticity of substitution, β is a positive distribution parameter, y_j is the nominal income of country j, and p_{ij} is the price of country i goods in country j, and $t_{ij} \geq 1$ is a 'iceberg' transportation cost from country j to country i.

By solving Equation (1), the nominal demand for the goods from country i by country j is derived as follows:

$$x_{ij} = (\frac{\beta_i p_i t_{ij}}{P_j})^{1-\sigma} y_j \quad \ldots (2)$$

where P_{ij} is the consumer price index in country j, defined as follows:

$$P_j = \left[\sum_i (\beta_i p_i t_{ij})^{1-\sigma} \right]^{1/(1-\sigma)} \quad \ldots (3)$$

The market clearing condition is given by the following equation:

$$y_i = \sum_j x_{ij} = \sum_j (\beta_i t_{ij} p_i / P_j)^{1-\sigma} y_j, \ \forall i \ \ldots(4)$$

Combining the above three equations and assuming symmetric transport cost ($t_{ij} = t_{ji}$), these equations can be dramatically simplified as follows:

$$x_{ij} = \frac{y_i y_j}{y^W} \left(\frac{t_{ij}}{P_i P_j}\right)^{1-\sigma} \quad \ldots(5)$$

where y^W is world income. P_j should satisfy the following equation:

$$P_j^{1-\sigma} = \sum_i \frac{y_i}{y^W} P^{\sigma-1} t_{ij}^{1-\sigma}, \ \forall j \quad \ldots(6)$$

Equation (5) is a form of the gravity equation. The trade volume (x_{ij}) increases as the product of GDP in country *j*(y_j) and GDP in country *i*(y_i) increases, and decreases as the distance (trade costs) between country *i* and *j* increases.

4.2 Equation and variables used in this paper

In the following estimations, I have tried as much as possible to use the same specifications as Rauch and Trindade (2002) in order for the results to be comparable. However, I have modified their specification in several ways in order to address some problems.

The basic specification of the gravity equation used in this paper is as follows.

$$
\begin{aligned}
V_{ijk} &= \alpha_k (GDP_i GDP_j)^{\beta_k} (PGNI_i PGNI_j)^{\gamma_k} \\
&\times DISTANCE^{\delta_k} REMOTE^{\varepsilon_k} \\
&\times \exp(\zeta_k ADJACENT + \eta_k EEC + \theta_k EFTA \qquad \ldots(7) \\
&+ \lambda_k LANGUAGE + \varphi_k COLOTIE + \psi_k CHINSHARE \\
&+ \tau_k LJPNSHARE + \omega_k PJPNSHARE) + u_{ijk}
\end{aligned}
$$

Two data sets have been used in this paper. One is the **basic country set.**, I have use a 62-country set which is almost identical to the 63-country data set used in Rauch and Trindade (2002), except for the exclusion of Libya.[3] The other data set is the **extended country set** consisting of 139 countries (Table 9), all the reporting countries covered by the PC-TAS plus Taiwan (see Appendix). Considering the selection bias, the larger the country set is the better. Thus, *i* takes 1 to 62 for the basic country set, and takes 1 to 139 for the extended country set.

The dependent variable, V_{ijk}, is the nominal value of imports of country *i* from *j*. *k* takes either 1 or 2, where *k*=1 denotes homogeneous goods, and *k*=2 denotes differentiated goods. The definitions of both goods are scripted in the Appendix.

Rauch and Trindade define V_{ijk} as the nominal value of bilateral trade (exports plus imports) between countries *i* and *j*. I decided to separate exports and imports so as not to lose

[3] Libya is excluded from the data set because various data for the country were not available in 2000.

any information by aggregating them. In Rauch and Trindade, k takes 1, 2, or 3, where k=1 denotes the organized exchange commodity group, k=2 denotes the reference-priced commodity group, and k=3 denotes the differentiated commodity group. The difference in the definition of k is discussed in the Appendix.

The description of the independent variables and the differences from Rauch and Trindade (2002) are as follows.

GDP denotes nominal GDP. Although Rauch and Trindade (2002) use nominal GNP, I use nominal GDP because of data availability;

PGNI denotes per capita nominal GNI (gross national income). Although Rauch and Trindade (2002) use per capita nominal GDP, I use nominal per capita GNI because of data availability;

DISTANCE equals the great circle distance between the capital cities of country i and j;

REMOTE equals the product of the weighted sum of country i's distances from all other countries in the sample and the same weighted sum for country j. The weights are the GDPs of other countries in the sample;

ADJACENT equals 1 if country i and j share a common land border, equals 0 otherwise;

EEC or EFTA equals 1 if both country *i* and *j* are a member of the EEC or EFTA, 0 otherwise;

LANGUAGE equals 1 if both country *i* and *j* share a common national language, equals 0 otherwise;

COLOTIE equals 1 if country *i* and *j* have a past colonial relationship, or both countries were the colonies of the same country, equals 0 otherwise;

CHINSHARE equals the product of the share of the ethnic Chinese population for country *i* and *j*;

LJPNSHARE equals the product of the share of the population of long-term Japanese stayers for country *i* and *j*.

PJPNSHARE equals the product of the share of the population of permanent Japanese stayers for country *i* and *j*.

A detailed description of the data is given in the Appendix. The summary statistics are presented in Table 1.

Table 1: SUMMARY STATISTICS FOR KEY VARIABLES 2000

		Mean	Standard Deviation	Minimum	Maximum
62 country-set	ln(GDP_iGDP_j)	51.04	2.258	49.9	59.11
(n = 3782)	ln(PGNI_iPGNI_j)	17.08	2.1788	10.2	21.06
	ln(DISTANCE)	8.7	0.851	5.153	9.899
	LANGUAGE	0.1	0.3	0	1
	COLOTIE	0.118	0.323	0	1
	CHINSHARE	0.003279	0.04326	0	0.939
	LJPNSHARE	0.000009984	0.0001621	0	0.005491
	PJPNSHARE	0.000003575	0.000042629	0	0.0009661
139 country-set	ln(GDP_iGDP_j)	47.22	3.282	36.46	59.11
(n = 19182)	ln(PGNI_iPGNI_j)	15.65	2.19	9.306	21.23
	ln(DISTANCE)	8.691	0.812	4.037	9.899
	LANGUAGE	0.111	0.314	0	1
	COLOTIE	0.1575	0.364	0	1
	CHINSHARE	0.0007397	0.019	0	0.939
	LJPNSHARE	0.000002466	0.00007321	0	0.005491
	PJPNSHARE	0.0000007354	0.000019	0	0.0009661

4.3 Methods of estimation

The value of bilateral trade is bounded by zero. Actually, about a half of bilateral trade values are zero for the basic country set. Therefore, the Ordinary Linear Square (OLS) estimator may be biased, and the Tobit estimator is generally used instead in this kind of truncated or censored sample estimation.

However, the Tobit estimator also cannot be used directly in the estimation of the gravity equation because the functional form of the gravity equation is usually log-linear, and ln(0) goes to -*Inf*.

To address this problem, Rauch and Trindade (2002) use the ``threshold Tobit'' estimator, a special Tobit estimator. I use the same threshold Tobit estimator in this paper. Details about the procedure are in Eaton and Tamura (1994).

Eaton and Tamura (1994) assumes that the goods of a_k thousand dollars are 'melted away' immediately after departing the exporter's port. In other words, the value of export goods that are going to be traded should be at least a_k thousand dollars to be traded. The Eaton and Tamura (1994) method estimates this threshold a_k, and thereafter follows the Tobit estimation. By this procedure, it is possible to avoid the 'ln(0)=-*Inf*' problem.

Thus, the basic equation estimated in this paper is as follows:

$$\begin{aligned} \ln(V_{ijk}) = \max[&\ln\alpha_k + \beta_k \ln(GDP_i GDP_j) + \gamma_k \ln(PGNI_i PGNI_j) \\ &+\delta_k \ln DISTANCE + \varepsilon_k \ln REMOTE \\ &+\zeta_k ADJACENT + \eta_k EEC + \theta_k EFTA \\ &+\lambda_k LANGUAGE + \varphi_k COLOTIE + \psi_k CHINSHARE \\ &+\tau_k LJPNSHARE + \omega_k PJPNSHARE) + u_{ijk}, \ln a_k] \end{aligned} \quad \ldots(8)$$

5. Results of Estimation

5.1 Effects of the ethnic Chinese networks in 1990

Table 2 shows the results of the main estimation by Rauch and Trindade (2002) on the

1990 data set under the conservative aggregation.

The main findings of Rauch and Trindade (2002) are as follows. In the first three columns of Table 2, the coefficients on *CHINSHARE* for all three goods (Org., Ref., and Dif. i.e., goods having organized exchanges, goods having referenced prices, and differentiated goods, respectively) are statistically significant at least at the 10% level. Moreover, the magnitude of coefficient on *CHINSHARE* is the largest for differentiated goods, and the smallest for goods that have organized exchanges. This result supports the search/matching view of international trade, i.e., ethnic Chinese networks are effective in overcoming the unnformativeness of goods, especially of differentiated goods, in international trade.

The latter three columns of Table 2 show the decomposition of the effect of ethnic Chinese networks in two categories. The dummy variable *TWO90ONE* equals one if both trading partners have a Chinese population of more than 1%. Thus, *CHINSHARE*(TWO90ONE)* captures the effect of ethnic Chinese networks in a pair of countries that have a large presence of ethnic Chinese, such as some East Asian countries. On the other hand, *CHINSHARE*(1-TWO90ONE)* captures the effect of ethnic Chinese networks for a pair of the countries that have a small presence of ethnic Chinese.

Almost all the coefficients in the latter three columns are not significantly different from those in the first three columns. Thus, support for the network/search view is largely retained even after decomposing the effect of ethnic Chinese networks.

Table 2: DEPENDENT VARIABLE: LOG OF 1990 BILATERAL TRADE IN ORGANISED EXCHANGE, REFERENCE PRICED, AND DIFFERENTIATED COMMODITIES (CONSERVATIVE AGGREGATION)

	Org.	Ref.	Dif.	Org.	Ref.	Dif.
Intercept	−45.295	−26.422	−19.805	−45.057	−25.804	−18.457
	(−3.601)	(−2.649)	(−2.570)	(−3.602)	(−2.645)	(−2.532)
Threshold(¥US thous.)	107.518 a	141.481 a	131.468 a	107.425 a	141.817 a	133.457 a
	(−14.155)	(−20.069)	(−22.530)	(−14.145)	(−20.112)	(−22.867)
ln(GNP_i GNP_j)(1990)	1.046 a	0.969 a	0.981 a	1.043 a	0.962 a	0.964 a
	(−0.035)	(−0.024)	(−0.024)	(−0.036)	(−0.025)	(−0.024)
ln(PGDP_i PGDP_j)(1990)	0.155 a	0.227 a	0.271 a	0.153 a	0.224 a	0.263 a
	(−0.039)	(−0.027)	(−0.026)	(−0.039)	(−0.026)	(−0.026)
ln(DISTANCE)	−1.23 a	−0.863 a	−0.677 a	−1.237 a	−0.881 a	−0.715 a
	(−0.108)	(−0.087)	(−0.088)	(−0.109)	(−0.087)	(−0.088)
ln(REMOTE)	2.148 a	0.989 a	0.518 a	2.142 a	0.974 a	0.487 a
	(−0.208)	(−0.159)	(−0.158)	(−0.208)	(−0.159)	(−0.156)
ADJACENT	0.818 b	0.921 a	1.038 a	0.811 b	0.902 a	0.997 a
	(−0.340)	(−0.286)	(−0.287)	(−0.340)	(−0.286)	(−0.286)
EEC	0.098	0.359 b	0.425 b	0.102	0.371 b	0.454 a
	(−0.225)	(−0.177)	(−0.169)	(−0.224)	(−0.176)	(−0.167)
EFTA	−0.264	0.303	0.489 b	−0.26	0.313	0.512 b
	(−0.440)	(−0.206)	(−0.231)	(−0.440)	(−0.206)	(−0.230)
LANGUAGE	0.903 b	1.142 a	0.316 b	0.913 a	1.166	0.368
	(−0.439)	(−0.341)	(−0.324)	(−0.440)	(−0.341)	(−0.323)
COLOTIE	0.303	0.472 a	0.934 a	0.298	0.458 a	0.903 a
	(−0.205)	(−0.153)	(−0.154)	(−0.205)	(−0.153)	(−0.153)
CHINSHARE	2.261 c	3.208 a	4.95 a			
	(−1.221)	(−0.639)	(−0.788)			
CHINSHARE*(1−TWO00ONE)				101.258	257.393 a	560.476 a
				(−99.561)	(−78.217)	(−111.286)
CHINSHARE*TWO00ONE				2.256 c	3.195 a	4.92 a
				(−1.222)	(−0.638)	(−0.775)
Log likelihood	−17813.8	−19086.4	−20763.5	−17813.8	−19084.7	−20752.9

(Source) Rauch and Trindade(2002), Table 5
Maximum likelihood estimation of threshold Tobit model.
Eicker-White standard errors in parentheses. Number of observations = 1710.
a) Significant at 1% level; b) Significant at 5% level; c) Significant at 10% level.

However, two new issues are revealed in the latter three columns. First, the coefficient on *CHINSHARE*(1-TWO90ONE)* for goods having organized exchanges is statistically insignificant even at the 10% level. The same coefficient on 1980 data sets (in Table 3 and 4 in Rauch and Trindade (2002)) is statistically significant at the 1% level. This might mean the declining importance of ethnic Chinese networks during the 1980s in the trading of goods that have organized exchanges.

Second, the coefficients on *CHINSHARE*(1-TWO90ONE)* are about 50 to 100 times larger than the coefficients on *CHINSHARE*(TWO90ONE)*. Rauch and Trindade (2002) explain this result as follows.

> [T]his diminishing marginal effects arises because ethnic Chinese communities becomes less cohesive as their population shares increase, but more likely it reflects the fact that countries with large ethnic Chinese populations shares also have large ethnic Chinese populations, diminishing the thoroughness with which any ethnic Chinese population increment is connected to the existing ethnic Chinese populations.(Rauch and Trindade 2002, p.125)

This 'diminishing marginal effects' of ethnic networks is examined in a later subsection.

5.2 **Effects of ethnic Chinese networks in 2000**

First, I estimated the equation that only contains the corresponding variables in Rauch and Trindade (2002) in order to test the effects of ethnic Chinese networks on international trade in 2000. Table 3 shows the results of the estimation using the data set in 2000 and the procedures mentioned above.

The coefficient on *CHINSHARE* for differentiated goods is larger than that for homogeneous goods, and statistically significant even at the 1% level. This result reinforces the arguments of Rauch and Trindade (2002).

On the other hand, the coefficient on *CHINSHARE* for homogeneous goods is significantly smaller than that for goods traded via organized exchanges in Table 2, and statistically insignificant even at the 10% level. Once the coefficients on *CHINSHARE* is decomposed into *CHINSHARE*(1-TWO00ONE)* and *CHINSHARE*TWO00ONE* [4], it becomes clear that only the coefficient of *CHINSHARE*TWO00ONE* for homogeneous goods is statistically insignificant, and significantly smaller than that in Table 2.

One possible explanation for this result is that the importance of ethnic Chinese networks for homogeneous goods significantly weakened in the 1990s. You will remember that in Table 2, the coefficient on *CHINSHARE*(1-TWO90ONE)* for goods that have organized exchanges is statistically insignificant even at the 10% level, while the same coefficient on the 1980 data sets is statistically significant at the 1% level[5]. Thus, this weakening effect of ethnic Chinese networks for homogeneous goods seems to be plausible.

---

[4] *TWO00ONE* is defined the same as is *TWO90ONE* on the year 2000 data set.

[5] In Table 3, the coefficient on *CHINSHARE*(1-TWO90ONE)* for homogeneous goods becomes significant again, but the classification used in this paper does not completely correspond to that used in Table 2.

Table 3: DEPENDENT VARIABLE: LOG OF 2000 BILATERAL TRADE IN HOMOGENEOUS AND DIFFERENTIATED COMMODITIES

	Hom.		Dif.		Hom.		Dif.	
Intercept	−51.56		−54.289		−51.063		−51.602	
	(−1.856)		(−2.049)		(−1.861)		(−2.029)	
Threshold(US thous.)	309.371	a	211.274	a	309.731	a	219.46	a
	(−23.550)		(−17.049)		(−23.575)		(−17.743)	
ln(GDP_i GDP_j)(2000)	0.9	a	1.031	a	0.897	a	1.015	a
	(−0.017)		(−0.019)		(−0.017)		(−0.019)	
ln(PGNI_i PGNI_j)(2000)	0.08	a	0.195	a	0.077	a	0.176	a
	(−0.015)		(−0.017)		(−0.015)		(−0.017)	
ln(DISTANCE)	−1.124	a	−1.068	a	−1.131	a	−1.1	a
	(−0.040)		(−0.045)		(−0.040)		(−0.045)	
ln(REMOTE)	1.341	a	0.989	a	1.328	a	0.915	a
	(−0.094)		(−0.104)		(−0.094)		(−0.103)	
ADJACENT	0.142		0.13		0.143		0.136	
	(−0.128)		(−0.144)		(−0.128)		(−0.144)	
EEC	0.378	b	0.025		0.387	b	0.076	
	(−0.153)		(−0.147)		(−0.153)		(−0.149)	
EFTA	0.392	a	0.525	a	0.398	a	0.556	a
	(−0.143)		(−0.117)		(−0.143)		(−0.116)	
LANGUAGE	0.64	a	0.476	a	0.63	a	0.421	a
	(−0.094)		(−0.111)		(−0.094)		(−0.110)	
COLOTIE	0.091		0.149		0.095		0.167	
	(−0.097)		(−0.112)		(−0.097)		(−0.111)	
CHINSHARE	0.829		3.596	a				
	(−0.674)		(−0.787)					
CHINSHARE*(1−TWO00ONE)					86.139	b	436.256	a
					(−33.974)		(−49.110)	
CHINSHARE*TWO00ONE					0.841		3.661	a
					(−0.674)		(−0.784)	
Log likelihood	−41547.9		−41773.4		−41545.9		−41733.3	

Maximum likelihood estimation of threshold Tobit model.
Eicker-White standard errors in parentheses. Number of observations = 3782.
a) Significant at 1% level; b) Significant at 5% level; c) Significant at 10% level.

The coefficients on *LANGUAGE* and *COLOTIE*, which are related to the network/search view, are both positive, but only the coefficients on *LANAGUGE* are statistically significant. In Table 2 the coefficients on *LANGUAGE* tend to be significant for the goods that have organized exchanges, while the coefficients on *COLOTIE* tend to be significant for differentiated goods. This result can be interpreted as showing that former colonial ties became less important in the 1990s while the importance of a common language became more important in international trade.

The other coefficients presented in Table 3 are relatively similar to the corresponding coefficients in Table 2. The coefficients on the logarithms of the product of *GDP*s, *PGNI*s are positive, statistically significant, and exhibit very similar values to those coefficients in Table 2. The coefficients on the logarithm of *DISTANCE* are negative, as expected, statistically significant, and again exhibit very similar values to those coefficients in Table 2. The coefficients on the logarithms of *REMOTE* and estimated Threshold are positive and statistically significant. The coefficients on *ADJACENT* are positive, as expected, but not statistically significant even at the 10% level. The coefficients on *EEC* are positive, but statistically significant only for homogeneous goods. This result is the reverse of Table 2, in which only the coefficients for differentiated goods are statistically significant. The coefficients on *EFTA* are positive and statistically significant both for homogeneous and differentiated goods. In Table 2 the coefficients on *EFTA* are statistically significant only for differentiated goods.

5.3 Effects of ethnic Japanese networks in 2000

To compare the effect of Japanese and Chinese networks, I estimated Equation (8) using all variables. The results are shown in Table 4. The first two columns show the estimated coefficients on the basic 62-country data set. The last two columns show the estimated coefficients on the extended 139-country data set. For the ethnic Japanese variables, there is no *TWO00ONE* term because no country other than Japan has a Japanese population of more than 1% of the total population.

In the first two columns, the coefficient on *LJPNSHARE* for homogeneous goods is negative, and that for differentiated goods is positive. This seems to support the search/matching view of trade, but both coefficients are statistically insignificant even at the 10% level. This indicates that ethnic Japanese modern networks seems to have some effects on international trade, but it is not possible to confirm the effects statistically on the basic country set.

On the other hand, the coefficient on *PJPNSHARE* for homogeneous goods and that for differentiated goods are both negative and statistically insignificant even at the 10% level indicating that traditional networks of ethnic Japanese seem to have no effects on international trade.

In the last two columns, the coefficient on *LJPNSHARE* for homogeneous goods is positive, but statistically insignificant even at the 10% level. However, that for differentiated goods becomes positive and statistically significant at the 1% level. The magnitude of the coefficient is almost equal to that on *CHINSHARE*(1-TWO00ONE)*.

Table 4: DEPENDENT VARIABLE:LOG OF 2000 BILATERAL TRADE IN HOMOGENEOUS AND DIFFERENTIATED COMMODITIES

	Hom.	Dif.	Hom.	Dif.
Intercept	−51.67	−51.776	−34.566	−33.233
	(−1.891)	(−2.073)	(−1.154)	(−1.224)
Threshold(US thous.)	309.013 a	219.476 a	95.262 a	72.024 a
	(−23.534)	(−17.758)	(−2.615)	(−2.047)
ln(GDP_i GDP_j)(2000)	0.901 a	1.017 a	0.789 a	0.869 a
	(−0.017)	(−0.019)	(−0.009)	(−0.010)
ln(PGNI_i PGNI_j)(2000)	0.079 a	0.177 a	0.047 a	0.161 a
	(−0.015)	(−0.017)	(−0.009)	(−0.010)
ln(DISTANCE)	−1.134 a	−1.097 a	−1.043 a	−1.11 a
	(−0.040)	(−0.045)	(−0.025)	(−0.026)
ln(REMOTE)	1.35 a	0.919 a	0.677 a	0.305 a
	(−0.095)	(−0.105)	(−0.063)	(−0.067)
ADJACENT	0.136	0.14	0.797 a	0.62 a
	(−0.128)	(−0.144)	(−0.102)	(−0.104)
EEC	0.373 b	0.074	0.343 b	0.196
	(−0.154)	(−0.149)	(−0.178)	(−0.142)
EFTA	0.394	0.558	0.609	0.594
	(−0.143)	(−0.116)	(−0.128)	(−0.101)
LANGUAGE	0.626 a	0.421 a	0.5 a	0.457 a
	(−0.094)	(−0.110)	(−0.063)	(−0.066)
COLOTIE	0.091	0.166	0.515 a	0.463 a
	(−0.097)	(−0.111)	(−0.056)	(−0.059)
CHINSHARE*(1−TWO00ONE)	83.484 b	432.349 a	104.095 a	401.763 a
	(−33.986)	(−48.899)	−32.347	−40.632
CHINSHARE*TWO00ONE	0.808	3.655 a	2.133 a	4.946 a
	(−0.673)	(−0.783)	(−0.720)	(−0.884)
LJPNSHARE	−119.465	167.172	90.951	396.772 a
	(−155.148)	(−143.805)	(−133.418)	(−131.059)
PJPNSHARE	−830.586	−920.728	808.917	933.697
	(−601.663)	(−583.260)	(−582.204)	(−585.630)
Log likelihood	−41544.6	−41732.4	−105532	−103830
Number of obs.	3782	3782	19182	19182

Maximum likelihood estimation of threshold Tobit model.
Eicker-White standard errors in parentheses.
a) Significant at 1% level; b) Significant at 5% level; c) Significant at 10% level.

Thus, on the extended country set, 'modern' networks of ethnic Japanese seem to have some positive effects on the trade of differentiated goods. Moreover, the coefficient for differentiated goods is significantly larger than that for homogeneous goods. This result supports the search/matching view of international trade.

As for the coefficients on *PJPNSHARE*, that for homogeneous goods and for differentiated goods are both positive, but statistically insignificant even at the 10% level. In addition, the difference between the coefficient for homogeneous goods and that for differentiated goods is not significant. Although 'traditional' networks of ethnic Japanese seem to have some effects on international trade, it is not statistically significant.

Looking at the ethnic Chinese variables, the coefficients are largely consistent with the former estimations. In the first two columns, the coefficients on *CHINSHARE*(1-TWO00ONE)* and *CHINSHARE*TWO00ONE* are basically compatible with the last two columns of Table 3. Thus the coefficients of the ethnic Chinese variables are not affected by adding ethnic Japanese variables.

On the other hand, there are some differences in the coefficients on ethnic Chinese variables between the basic country set and the extended country set. The coefficients on *CHINSHARE*TWO00ONE* for homogeneous goods are significantly larger on the extended country set than the basic country set. This seems to be the result of the inclusion of several small countries, such as Costa Rica, Mauritius, Panama, Surinam, and Trinidad Tobago. These countries have a high ratio of ethnic Chinese in their population. Thus, it can be concluded that ethnic Chinese networks still play important roles in the trade of homogeneous goods between small countries that have higher ratios of ethnic Chinese in the national population.

As for the coefficients on *LANGUAGE* and *COLOTIE*, that on *LANGUAGE* is not significantly different on either the basic or extended country sets, while the coefficient on *COLOTIE* is significantly larger on the extended country set than the basic country set. This result is understandable considering that the extended country set includes smaller countries which seem to have a stronger dependence on colonial ties.

Regarding the other coefficients, there are some differences in the coefficients between the basic and extended country set. The coefficients on the logarithms of *REMOTE* are significantly smaller on the extended country set than the basic country set. On the other hand, the coefficients on *ADJACENT* are significantly larger on the extended country set than the basic country set.

5.4 Checking robustness

5.4.1 Diminishing returns and emigrant effect

The large difference in the coefficients of *CHINSHARE*(1-TWO00ONE)* and *CHINSHARE*TWO00ONE* means that the effects of ethnic networks have diminished as the networks have become larger. To check these diminishing effects, Rauch and Trindade (2002) use other variables proxying the extensiveness of the ethnic networks instead of *CHINSHARE*. They use the product of the absolute number of ethnic Chinese in both trading partners, *CHINPOP*, as a measure of the extensiveness of ethnic Chinese networks. They also include the quadratic term of *CHINPOP*. If ethnic networks exhibit diminishing effects on international trade, the coefficient on the *CHINPOP* should be positive, and that on $CHINPOP^2$ should be negative.

Rauch and Trindade (2002) also separate country pairs into two categories: pairs that include China and pairs that do not include China, because China has a Chinese population more than fifty times larger than any other country. Therefore they add the dummy variable *CHINA*, which is 1 if the country pair includes China, and 0 otherwise.

This separation has another benefit. If the effects of ethnic Chinese networks on international trade are significantly larger for the country pairs including China than for other country pairs, then this effect is regarded as the emigrant effect studied by Gould (1994), rather than general network effects.

Rauch and Trindade (2002) proved that the effects of ethnic Chinese networks are diminishing, and the effects are statistically significant even if the country pairs which include China are separated (See Table 7 in Rauch and Trindade (2002)). Therefore the effects of ethnic Chinese networks on international trade are interpreted as general network effects, rather than emigrant effects.

In this paper I have taken the logarithms of the product of the absolute number of ethnic Chinese (Japanese) in both trading partners to test the diminishing effects of ethnic networks, because it is not possible to implement the Rauch and Trindade (2002) method due to the insufficient accuracy of computation[6]. If the log specification has a higher statistical

[6] This problem is explained in the Conclusions.

significance than the liner specification used in Table 4, it means that the effects of the networks on international trade are diminishing with their increasing size.

At the same time, I have separated the country pairs into two categories, the pairs include China (Japan), and other pairs. I have introduced the dummy variable *CHINA* (*JAPAN*) if the country pair includes China (Japan). The results are presented in Table 5.

The coefficients on *ln(CHINPOP)*(1-CHINA)* and *ln(CHINPOP)*CHINA* are both statistically significant at the 1% level, and statistical significance is largely improved over the coefficients on *CHINSHARE*(1-TWO00ONE)* and *CHINSHARE*TWO00ONE* in Table 4. Moreover, the coefficients on *ln(CHINPOP)*(1-CHINA)* and those on *ln(CHINPOP)*CHINA* are not significantly different. In Table 4 the difference between *CHINSHARE*(1-TWO00ONE)* and *CHINSHARE*TWO00ONE* is huge. This means that Chinese networks exhibit diminishing returns to the size of the networks, and seems to fit to the log-specification.

Furthermore, both the coefficients on *ln(CHINPOP)*(1-CHINA)* and on *ln(CHINPOP)*CHINA* for differentiated goods are always significantly larger than those coefficients for homogeneous goods. Thus, the validity of the search/matching view of international trade remains for the networks of ethnic Chinese in 2000.

Regarding the coefficients on *ln(LJPNPOP)*(1-JAPAN)* and *ln(LJPNPOP)*JAPAN*, the statistical significance is generally improved over the significance of the coefficients on *LJPNSHARE* in Table 4. In Table 5 the coefficients on *ln(LJPNPOP)*(1-JAPAN)* are always statistically significant on both basic and extended country set, regardless of the type of goods traded. The coefficients on *ln(LJPNPOP)*JAPAN* are statistically significant

Table 5: DEPENDENT VARIABLE:LOG OF 2000 BILATERAL TRADE IN HOMOGENEOUS AND DIFFERENTIATED COMMODITIES

	Hom.	Dif.	Hom.	Dif.
Intercept	−43.755	−24.185	−24.469	−17.847
	(−2.413)	(−2.506)	(−1.433)	(−1.521)
Threshold(US thous.)	306.947 a	237.279 a	95.924 a	72.432 a
	(−23.420)	(−19.665)	(−2.631)	(−2.056)
ln(GDP_i GDP_j)(2000)	0.805 a	0.68 a	0.67 a	0.679 a
	(−0.023)	(−0.025)	(−0.013)	(−0.014)
ln(PGNI_i PGNI_j)(2000)	0.098 a	0.196 a	0.046 a	0.171 a
	(−0.015)	(−0.017)	(−0.010)	(−0.010)
ln(DISTANCE)	−1.148 a	−1.147 a	−1.081 a	−1.171 a
	(−0.040)	(−0.042)	(−0.024)	(−0.026)
ln(REMOTE)	1.108 a	0.123	0.397 a	−0.114
	(−0.103)	(−0.105)	(−0.067)	(−0.071)
ADJACENT	0.218 c	0.28 b	0.731 a	0.548 a
	(−0.131)	(−0.140)	(−0.102)	(−0.103)
EEC	0.059	−0.597 a	0.167	−0.18
	(−0.157)	(−0.200)	(−0.184)	(−0.162)
EFTA	0.294 b	0.39 a	0.488 a	0.406 a
	(−0.145)	(−0.121)	(−0.131)	(−0.106)
LANGUAGE	0.579 a	0.397 a	0.468 a	0.416 a
	(−0.094)	(−0.101)	(−0.062)	(−0.064)
COLOTIE	0.118	0.133	0.447 a	0.382 a
	(−0.095)	(−0.105)	(−0.056)	(−0.058)
ln(CHINPOP)*(1−CHINA)	0.042 a	0.079 a	0.019 a	0.028 a
	(−0.005)	(−0.006)	−0.002	−0.002
ln(CHINPOP)*CHINA	(0.048) a	(0.083) a	0.031 a	0.053 a
	(−0.007)	(−0.008)	−0.004	−0.005
ln(LJPNPOP)*(1−JAPAN)	(0.048) a	(0.203) a	0.063 a	0.115 a
	(−0.012)	(−0.016)	−0.006	−0.008
ln(LJPNPOP)*JAPAN	0.04	0.246 a	−0.009	0.06 a
	(−0.036)	(−0.030)	−0.009	−0.012
ln(PJPNPOP)*(1−JAPAN)	−0.034 a	−0.039 a	0.014 a	0.01 a
	(−0.005)	(−0.006)	(−0.004)	(−0.004)
ln(PJPNPOP)*JAPAN	−0.018	−0.072 b	0.056 a	0.031 a
	(−0.041)	(−0.037)	(−0.010)	(−0.012)
Log lilklihood	−41500	−41581.2	−105450.3	−103717.3
Number of obs.	3782	3782	19182	19182

Maximum likelihood estimation of threshold Tobit model.
Eicker-White standard errors in parentheses.
a) Significant at 1% level; b) Significant at 5% level; c) Significant at 10% level.

for differentiated goods on both the basic and extended country sets. This result shows that the 'modern' networks of ethnic Japanese exhibit diminishing returns to network size.

In addition, the coefficients on both *ln(LJPNPOP)*(1-JAPAN)* and *ln(LJPNPOP)*JAPAN* for differentiated goods are significantly larger than those coefficients for homogeneous goods on both the basic and extended country sets. This result supports the search/matching view of international trade, and shows that the effects of 'modern' Japanese networks are not emigrant effects.

For the coefficients on *ln(PJPNPOP)*(1-JAPAN)* and *ln(PJPNPOP)*JAPAN*, the statistical significance is generally improved over that of the coefficients on *PJPNSHARE* in Table 4. However, in the first two columns, the coefficients on *ln(LJPNPOP)*(1-JAPAN)* and *ln(LJPNPOP)*JAPAN* are all negative regardless of the type of goods. In the last two columns, all those coefficients become positive and statistically significant. But the coefficients for homogeneous goods are higher than those for differentiated goods. This does not fit the prediction of the search/matching view of international trade. Moreover, the coefficients on *ln(LJPNPOP)*JAPAN* are higher than those on *ln(LJPNPOP)*(1-JAPAN)*. This result suggests that the effects of 'traditional' Japanese networks are emigrant effects.

5.4.2 Overall robustness with full-set country dummy variables

Anderson and Wincoop (2002) argue critically that the majority of the gravity equations estimated in various empirical works do not correspond to the gravity equations derived theoretically. They point out that the often-used term, 'remoteness' index is too ad hoc and has nothing to do with the theory behind the gravity equation. They propose instead the term 'multilateral trade resistance'. Multilateral trade resistance is defined as Equation (3) in section 3 of this paper.

Rauch and Trindade (2002) use an ad hoc remoteness index criticized by Anderson and Wincoop. Thus, to test the robustness of the results, this problem needs to be addressed.

Redding and Venables (2004) estimated the gravity equation by adding a full set of country dummies for both importing and exporting countries. This makes possible an equally accurate estimation to that using multilateral trade resistance without knowing the price data of each country. This greatly reduces the amount of work on the estimation, especially for one involving more than two countries.

In the following estimation I have adopted the same approach as Redding and Venables to test the robustness of the estimated results. More precisely, I have dropped all country-specific terms from the equation estimated in the previous subsection (the log-specification), and include the full set of country dummies.

The following equation is estimated[7].

$$\begin{aligned}\ln(V_{ijk} + a_k) = \max[&\ln\alpha_k + T_M D_M + T_X D_X \\ &+\delta_k \ln DISTANCE + \varepsilon_k \ln REMOTE \\ &+\zeta_k ADJACENT + \eta_k EEC + \theta_k EFTA \qquad \ldots(9) \\ &+\lambda_k LANGUAGE + \varphi_k COLOTIE + \psi_k CHINSHARE\end{aligned}$$

[7] *ln CHINPOP*, *ln LJPNPOP*, and *ln PJPNPOP* are approximated by *ln CHINPOP+1*, *ln LJPNPOP+1*, and *ln PJPNPOP+1*, respectively.

$$+\tau_k LJPNSHARE + \omega_k PJPNSHARE) + u_{ijk}, \ln a_k]$$

where D_M and D_X are the matrices of the dummy variables for import countries and export countries.

However, it is not possible to conduct the threshold Tobit estimation on Equation (9) due to computational limitation[8]. Therefore I conducted the ordinary Tobit estimation on Equation (9) using Thresholds a_k in Table 5. The results of the estimation are presented in Table 6.

Almost all the coefficients are statistically significant at the 1% level. The only exception is the coefficients on *ln LJPNPOP* for the basic country set, which are not obtained, because of the multicollinearity with full-set country dummies. The coefficients on *ln CHINPOP* are always positive and statistically significant at the 1% level. In addition, the coefficient for differentiated goods is bigger than that for homogeneous goods, although the difference is small on the basic data set. Thus the effects of ethnic Chinese networks on international trade are confirmed.

As for the effects of ethnic Japanese networks, the coefficients on *ln LJPNPOP* on the extended data set are positive and statistically significant. Although the coefficient for differentiated goods is larger than that for homogeneous goods, the difference is not statistically significant. Therefore the effects of 'modern' Japanese networks on international trade are only partly confirmed.

---

[8] the problem is detailed in the Conclusions.

Table 6: DEPENDENT VARIABLE:LOG OF 2000 BILATERAL TRADE IN HOMOGENEOUS AND DIFFERENTIATED COMMODITIES

	Hom.	Dif.	Hom.	Dif.
Intercept	15.43 a	14.768 a	9.432 a	8.489 a
	(−0.258)	(−0.231)	(−0.070)	(−0.071)
Threshold(US thous.)	306.947	237.239	95.262	72.024
ln(DISTANCE)	(−1.114) a	(−1.040) a	(−0.902) a	(−0.960) a
	−0.035	−0.033	−0.017	−0.017
ADJACENT	(0.225) a	(0.372) a	(0.838) a	(0.726) a
	−0.13	−0.123	−0.083	−0.083
EEC	(−0.985) a	(−0.947) a	(0.164) a	(0.280) a
	−0.289	−0.273	−0.267	−0.267
EFTA	(0.663) a	(0.410) a	(1.241) a	(1.229) a
	−0.162	−0.153	−0.174	−0.174
LANGUAGE	(0.496) a	(0.563) a	(0.248) a	(0.250) a
	−0.084	−0.079	−0.045	−0.045
COLOTIE	(0.350) a	(0.341) a	(0.512) a	(0.435) a
	−0.08	−0.075	−0.04	−0.04
ln(CHINPOP)	(0.082) a	(0.091) a	(0.051) a	(0.060) a
	−0.019	−0.018	−0.003	−0.003
ln(LJPNPOP)	n.a.	n.a	(0.127) a	(0.132) a
			−0.012	−0.012
ln(PJPNPOP)	(0.027) a	(0.045) a	(0.124) a	(0.134) a
	−0.013	−0.012	−0.004	−0.004
Log likelihood	−5973	−5759	−33507	−33510
Number of obs.	3782	3782	19182	19182

*Thresholds are the estimated values in Table 5.
Maximum likelihood estimation of threshold Tobit model.
Eicker-White standard errors in parentheses.
a) Significant at 1% level; b) Significant at 5% level; c) Significant at 10% level.

The coefficients on *ln PJPNPOP* are also positive and statistically significant at the 1% level. Additionally, the coefficient on differentiated goods is larger than that for homogeneous goods, although the difference on the basic data set is not statistically significant. It seems from this result that 'traditional' networks of ethnic Japanese have significant effects on international trade.

However, the results presented above are suspected of being affected by the multicollinearity with full-set country dummies, which makes *ln LJPNPOP* on the basic country set indeterminate. The number of long-term Japanese stayers in each country in the basic country set does not include any zeros. So *ln LJPNPOP* can be decomposed into the linear combinations of full-set country dummies weighted by the logarithms of the number of long-term Japanese stayers in each country. Thus, *ln LJPNPOP* exhibits perfect multicollinearity with full-set country dummies, and unobtainable.

In the case of *ln CHINPOP* and *ln PJPNPOP*, the corresponding Chinese population and permanent Japanese residents in each country contains some zeros. So the multicollinearity with full-set country dummies is not perfect, and the coefficients on those variables are obtainable. However, those coefficients are suspected of being affected by the very high correlation with the linear combination of full-set country dummies.

Thus, non-log specification should have been used in this robustness test with full-set country dummies. But it is not possible because of other computational difficulties, mentioned in the last section.

5.5 Quantitative importance of network effects

In addition to the statistical significance of networks effects on international trade, I also have calculated the quantitative importance of the effects by following Rauch and Trindade (2002). Table 7 shows the percentage of increase in bilateral trade attributable to ethnic Chinese networks and colonial ties in 1980 and 1990 calculated on the conservative aggregation. It shows that ethnic Chinese networks have quantitatively important effects on international trade, especially for bilateral trade between countries that have a Chinese population of more than one percent of total population.

Table 7: PERCENTAGE INCREASE IN BILATERAL TRADE ATTRIBUTABLE TO ETHNIC CHINESE NETWORKS AND COLONIAL TIES (Conservative Aggregation)

		Org.	Ref.	Dif.
1980	CHINSHARE*(1-TWO00ONE)	3.8	4.5	6.2
	CHINSHARE*TWO00ONE	88.8	128.3	177.8
	COLOTIE	9	13.7	18.6
1990	CHINSHARE*(1-TWO00ONE)	1	2.5	5.5
	CHINSHARE*TWO00ONE	23.8	35.4	59.2
	COLOTIE	4.4	6.8	13.8

(source) Column 1-3 of Table 9 in Rauch and Trindade(2002)

Table 8 shows the percentage of increase in bilateral trade in 2000 attributed to ethnic Chinese networks, 'modern' and 'traditional' networks of ethnic Japanese, common language effects and colonial ties. The first two columns show the percentage of increase calculated on the basic country set. The last two columns show the percentage of increase calculated on the extended country set. Only statistically significant variables are presented in the table.

Table 8: PERCENTAGE INCREASE IN BILATERAL TRADE ATTRIBUTED TO NETWORK EFFECTS (2000)

	Hom.	Dif.	Hom.	Dif.
CHINSHARE*(1-TWO00ONE)	1.2	6.3	0.6	2.5
CHINSHARE*TWO00ONE	–	62.1	12.6	31.6
LJPNSHARE	–	–	–	0.1
PJPNSHARE	–	–	–	–
LANGUAGE	6.5	4.3	6.5	5.8
COLOTIE	–	–	9.6	8.5
Number of Obs.	3782	3782	19182	19182

The quantitative importance of ethnic Chinese networks in 2000 seems to be almost the same as it was in 1990. However, the importance of these networks is smaller on the extended country set.

Regarding ethnic Japanese networks, only the 'modern' networks of ethnic Japanese have statistically significant effects for differentiated goods, on the extended country set. In addition, the percentage of increase in bilateral trade attributable to these networks is only 0.1%. It seems that ethnic Japanese networks have no quantitative importance in international trade.

This result is plausible because ethnic Japanese are a tiny percentage of the population in the host countries. Ethnic Chinese make up a larger percentage of the population in host countries than do ethnic Japanese. Table 1 shows that the mean value of *CHINSHARE* is several hundred times larger than that for *LJPNSHARE*, and about one thousand times larger than *PJPNSHARE*.

However, this does not mean the ethnic Japanese networks have no importance in international trade, because the quantity presented in Table 8 is calculated at the mean values of all variables. Thus, for some country pairs, ethnic Japanese networks have a much higher importance than the quantity presented in Table 8. For instance, in country pairs that include Japan, the value of *LJPNSHARE* is 68 times larger than the mean value of *LJPNSHARE* on the extended country set. Thus, the quantitative importance of ethnic Japanese networks is not negligible, at least for the country pairs that includieJapan.

6. Conclusions

In this paper I re-examined the statistical significance of ethnic Chinese networks on international trade, presented by Rauch and Trindade (2002), by using a newer and more extensive data set. In addition, I tested the statistical significance of 'modern' and 'traditional' ethnic Japanese networks on international trade. The main findings of this paper are as follows.

6.1 Main findings

First, the networks of ethnic Chinese have statistically significant effects on the volume of bilateral trade, except for homogeneous goods on the basic country set. Moreover, ethnic

Chinese networks increase the trade of differentiated goods more than that of homogeneous goods. Thus, both (null) Hypothesis 1a and 1b are rejected. This result is another supporting example to the network/search view of international trade.

Second, the 'modern' networks of ethnic Japanese have statistically significant effects on the volume of bilateral trade only for differentiated goods on the extend country set. This result is understandable considering that 'modern' Japanese networks are driven by multinational Japanese companies, and they trade in mainly differentiated goods. Thus, Hypothesis 2a is partly rejected. Beyond this, in almost all cases the coefficients on differentiated goods are significantly larger than those on homogeneous goods. Therefore, Hypothesis 2b is generally rejected. It can be concluded from this result that 'modern' networks of ethnic Japanese increase the volume of trade in differentiated goods, but the effects of Japanese networks are not as robust as those of ethnic Chinese networks.

Third, the 'traditional' networks of ethnic Japanese generally do not have statistically significant effects on the volume of bilateral trade. Therefore, Hypothesis 3a is not rejected in most cases. In some settings, the 'traditional' networks of ethnic Japanese have some positive and statistically significant effect on homogeneous goods in country pairs that include Japan. In this case, Hypothesis 3a is rejected and 3b is not rejected. This type of network effect seems to be an emigrant effect.

Fourth, combining the first and second findings, the effects of ethnic networks on international trade is not unique to the ethnic Chinese. Because of their large presence in many host countries, the networks of ethnic Chinese may have the biggest effects on international trade. However, these network effects on international trade can be generalized beyond the ethnic Chinese networks. This result adds further strong support to the search/matching view of international trade.

Fifth, combining the second and third findings, it can be concluded that the 'cohesiveness' of ethnic networks seems to matter for international trade. Because the 'modern' networks of ethnic Japanese have higher cohesiveness than the 'traditional' networks, the former have stronger effects on international trade than the latter. This result has some policy implications. The policies that keep overseas nationals united may enhance the cohesiveness of ethnic networks, and increase the volume of international trade through those networks.

6.2 Remaining issues

6.2.1 Refinement of data set

In this paper I have endeavored to replicate the estimation of Rauch and Trindade (2002) as much as possible in order for the results to be comparable. However, two main differences in the data set should be addressed in future research.

The biggest difference between the data set used in this paper and that in Rauch and Trindade (2002) is the classification of trading goods. Rauch and Trindade (2002) classified goods into three categories (i.e., goods traded via organized exchanges, goods having referenced prices, and differentiated goods) at the SITC 4-digit level. In this paper goods are classified into two categories, homogeneous goods and differentiated goods, at the SITC 1-digit level. To make the result fully comparable, the same classification should be used.

The second difference between the data set used in this paper and that used in Rauch and Trindade (2002) is the variable *LANGUAGE*. In Rauch and Trindade (2002), *LANGUAGE* is constructed by considering the proportion of the population that uses each language within a country. In this paper, *LANGUAGE* is constructed by only considering the most commonly used language in a country. Clearly the former variable represents more accurately the real world situation.

6.2.2 Computational problems

Through the different settings of estimations in this paper, the effects of the ethic Chinese on international trade is fairly robust. However, the effects of the ethic Japanese on international trade vary depending on the specifications of estimation. One reason is that the presence of ethic Japanese in host countries is generally so small that true trade enhancing effects are not easily detected.

Another reason of the fragility of Japanese coefficients is related to a more general computational problem. The population of Japanese in Japan is 5 million times larger than the population of Japanese in Sudan. This leads to a huge variance in some variables. For instance, in the basic country set, the maximum value of *LJPNPOP* is almost 30 billion times larger than the minimum value of *LJPNPOP*. This extremely large variation in some variables causes computational inaccuracy due to rounding and makes some coefficients indeterminant. This problem needs to be addressed in future research.

Another computational problem is the extensive computational requirement of the Newton iteration for the threshold Tobit estimation. It takes a very long time to estimate some models, especially when the model involves full-set country dummies. More computationally efficient algorithms will need to be implemented in future research.

6.3 Future research topics

Although research on ethnic Japanese networks has some important implication, their small representation in host countries make it difficult to conduct further detailed econometric studies. On the other hand, the networks of ethnic Chinese are easy to analyze in econometric studies because of their large presence in host countries. In this paper, the robustness of the effect of ethnic Chinese networks stood out.

Thus, conducting further research on the effects of ethnic Chinese network on international trade seems to be a reasonable way to proceed. For instance, estimating the effect of ethnic Chinese networks by the home province of immigrants would be an important research topic for reinforcing the network/search view of international trade.

Table 9: COUNTRIES, CHINESE POPULATION(CHINPOP), LONG-TERM JAPANESE STAYERS(LJPNPOP) AND PERMANENT JAPANESE RESIDENTS(PJPNPOP) IN 2000

	Country	CHINPOP	LJPNPOP	PJPNPOP
1	ALBANIA	0	0	0
2	ALGERIA*	2,000	62	10
3	ARGENTINA*	30,000	741	11,063
4	ARMENIA	7,317	0	0
5	AUSTRALIA*	454,000	21,614	16,813
6	AUSTRIA*	20,000	1,247	579
7	AZERBAIJAN	18,925	55	0
8	BAHAMAS	200	31	7
9	BAHRAIN	48	190	7
10	BANGLADESH	700	380	35
11	BARBADOS	50	7	9
12	BELARUS	23,524	17	1
13	BELGIUM*	30,000	4,936	0
14	BELIZE	1,500	11	11
15	BENIN	0	13	0
16	BHUTAN	0	52	0
17	BOLIVIA*	12,000	300	2,345
18	BOTSWANA	40	49	3
19	BRAZIL*	100,000	2,674	72,644
20	BULGARIA	25	122	13
21	BURUNDI	0	0	0
22	CAMEROON	50	40	0
23	CANADA*	910,000	13,580	20,486
24	CAPEVERDE	0	14	0
25	CHILE*	5,000	688	420
26	CHINA*	1,160,200,740	20,061	206
27	COLOMBIA*	7,000	398	994
28	COMOROS	0	0	0
29	COSTARICA	63,000	357	0
30	COTEDIVOIRE	200	182	0
31	CROATIA	0	32	25
32	CYPRUS	720	15	3
33	CZECHREP	100	420	42
34	DENMARK*	6,000	339	621
35	DOMINICA	0	3	0
36	ECUADOR*	20,000	230	181
37	EGYPT*	110	735	177
38	ELSALVADOR	1,500	143	49
39	ESTONIA	3,220	20	0
40	ETHIOPIA*	100	130	0
41	FINLAND*	1,000	321	424
42	FRANCE*	300,000	20,632	4,942
43	GABON	100	16	6
44	GAMBIA	0	7	0
45	GEORGIA	12,372	4	0
46	GERMANY*	111,000	21,237	3,784
47	GHANA*	500	242	13

(Source)See Appendix.
* is a country included in the 62 country set.

Table 9: COUNTRIES, CHINESE POPULATION(CHINPOP), LONG-TERM JAPANESE STAYERS(LJPNPOP) AND PERMANENT JAPANESE RESIDENTS(PJPNPOP) IN 2000 (CONT.)

	Country	CHINPOP	LJPNPOP	PJPNPOP
48	GREECE*	229	186	434
49	GRENADA	0	4	0
50	GUATEMALA	14,000	263	75
51	GUINEA	0	30	0
52	HONDURAS	1,500	160	55
53	HONGKONG*	6,331,750	25,363	460
54	HUNGARY*	10,000	767	72
55	ICELAND*	100	37	9
56	INDIA*	135,000	1,937	98
57	INDONESIA*	7,310,000	11,586	668
58	IRAN(ISLM.R)*	200	238	197
59	IRELAND*	5,000	527	225
60	ISRAEL*	225	364	254
61	ITALY*	30,000	6,549	1,448
62	JAMAICA	25,000	116	24
63	JAPAN*	170,000	126,925,843	126,925,843
64	JORDAN	200	216	27
65	KAZAKSTAN	35,408	113	0
66	KENYA*	190	735	0
67	KIRIBATI	0	33	0
68	KOREAREP.*	30,000	15,751	695
69	KUWAIT*	200	116	45
70	KYRGYZSTAN	11,556	38	0
71	LATVIA	5,577	11	0
72	LEBANON	12	47	25
73	LITHUANIA	8,243	25	2
74	LUXEMBOURG	6,500	366	0
75	MACEDONIA,TFYR	0	9	0
76	MADAGASCAR	30,000	117	4
77	MALAYSIA*	5,280,000	11,024	601
78	MALDIVES	0	109	0
79	MALI	0	20	0
80	MALTA	15	19	19
81	MAURITIUS	40,000	29	16
82	MEXICO*	30,000	2,588	1,570
83	MOLDOVAREP.	10,059	0	0
84	MONGOLIA	4,000	259	0
85	MOROCCO*	50	266	0
86	MOZAMBIQUE	700	60	0
87	NEPAL	20,348	408	0
88	NETHERLANDS*	80,000	5,722	759
89	NEWZEALAND*	35,000	4,077	3,703
90	NIGER*	20	80	0
91	NIGERIA*	2,000	96	24
92	NORWAY*	1,000	259	295
93	OMAN	80	100	15
94	PAKISTAN*	3,600	547	263

(Source)See Appendix.
* is a country included in the 62 country set.

Table 9: COUNTRIES, CHINESE POPULATION(CHINPOP), LONG-TERM JAPANESE STAYERS(LJPNPOP) AND PERMANENT JAPANESE RESIDENTS(PJPNPOP) IN 2000 (CONT.)

	Country	CHINPOP	LJPNPOP	PJPNPOP
95	PANAMA	150,000	431	0
96	PAPUA N.GUIN	10,000	228	0
97	PARAGUAY*	10,000	356	3,559
98	PERU*	60,000	588	1,222
99	PHILIPPINES*	2,200,000	7,980	1,247
100	POLAND*	200	531	123
101	PORTUGAL*	10,000	535	0
102	QATAR	0	112	1
103	ROMANIA	35	212	0
104	RUSSIAN FED	342,236	1,446	38
105	ST.KITTS NEV	0	0	0
106	S.VINCENT-GR	0	1	3
107	SAUDI ARABIA*	45,000	819	27
108	SENEGAL	0	140	0
109	SINGAPORE*	2,291,100	22,074	989
110	SLOVAKIA	0	66	2
111	SLOVENIA	0	24	10
112	SOUTH AFRICA*	30,000	1,085	125
113	SPAIN*	35,000	3,717	966
114	SRI LANKA	3,500	836	32
115	SUDAN*	45	23	12
116	SURINAME	13,000	32	1
117	SWAZILAND	90	7	0
118	SWEDEN*	20,000	728	1,414
119	SWITZ.LIECHT*	13,286	2,632	3,062
120	SYRIA A. R.	0	191	6
121	TAIWAN (POC)*	21,831,460	13,613	428
122	TAJIKISTAN	14,561	2	0
123	TANZANIA, U.R	600	296	2
124	THAILAND*	6,100,000	20,405	749
125	TOGO	30	5	0
126	TONGA	20	66	0
127	TRINIDAD TBG	20,000	39	2
128	TUNISIA*	0	139	0
129	TURKEY*	60,000	788	242
130	TURKMENISTAN	12,426	34	0
131	UGANDA	100	110	3
132	UNTD KINGDOM*	250,000	43,646	9,468
133	USA,PR,USVI*	2,000,000	188,360	109,608
134	URUGUAY*	300	89	267
135	VANUATU	0	70	0
136	VENEZUELA*	50,000	334	361
137	YUGOSLAVIA*	0	87	16
138	ZAMBIA	150	226	0
139	ZIMBABWE	300	196	22

(Source)See Appendix.
* is a country included in the 62 country set.

Appendix

Trade data

The trade data used in this paper was constructed from The Personal Computer Trade Analysis System (PC-TAS) published by the International Trade Center, UNCTAD/WTO. The PC-TAS is derived from the United Nations COMTRADE, which covers over 90% of world trade. Although the PC-TAS is a subset of the COMTRADE, it censored only the transactions of less than US$5,000 from the COMTRADE. The PC-TAS contains both export and import data for the latest five years, covering about 200 countries up to the Standard International Trade Classification (SITC rev.3) 5-digit level.

To construct bilateral trade data, V_{ij}, I used the import data of country *i*, not the export data of country *j*, because the import data should identify trade partners more accurately than the export data, especially when goods are traded through a third country.

The trade data of Taiwan as a reporting country is not available in the PC-TAS. Therefore I constructed the data on exports to Taiwan not from the import data of that country, but from the export data of each country exporting to Taiwan.

Number of overseas Japanese and Chinese

The data on the ethnic Japanese population overseas was derived from *Zairyu Hojin Tokei* (Statistics on overseas Japanese), the Ministry of Foreign Affairs, Japan. This data covers about 200 countries and regions. The number of overseas Japanese is the sum of two criteria: *Cho-ki Taizai-sha* (long-term stayers) and *Eiju-sha* (permanent residents). The former is defined as persons who have stayed in a country for more than three months and are not permanent residents. The latter is defined as persons who have acquired permanent residency in a country.

Regarding the number of Japanese in Hong Kong, having been integrated into China after 1998, I split the number of Japanese in China into those in China and those in Hong Kong, according to the proportion of Japanese in each location in 1997 (last reported.)

The data on the ethnic Chinese population overseas was obtained from ``Distribution of the Overseas Chinese Population" on the website of the World Confederation of Institutes and Libraries for Chinese Overseas Studies (WCILCOS). This data has been compiled by the Institute of Overseas Chinese Studies, Jinan University and covers 141 countries and regions.

Regarding the number of ethnic Chinese in the former Soviet Union countries (FSUs), these are reported as the combined total for the 'FSUs' as a whole. Therefore I distributed a number to each FSU country proportional to the total population of the countries, as a proxy for the actual number.

Other data

Geographical distance between two countries is approximated by the distance between the capital cities of two countries. This is calculated from the latitude and longitude data of the capital cities using great circle distance[9].

GDP and population data of each country were derived from Data Query of the World Bank, which provides access to a part of the World Development Indicators (WDI) database.

Other geographical, social, and historical data (contiguity, language and colonial ties) were derived from the World Fact Book by the Central Intelligence Agency of the United States. It covers virtually all countries and regions in the world giving detailed social, economical and political data.

Classification of goods

Rauch and Trindade (2002) classify goods into three categories in line with Rauch (1999): goods traded on organized exchanges, goods that have reference prices, and all other commodities,. For instance, oil and gold are the goods traded on organized exchanges, like the Chicago Mercantile Exchange. Goods that have reference prices are, for example, some chemical products that are not traded on organized exchanges but are highly standardized and have 'reference prices' in industry magazines or newspapers. The third category covers

[9] I wrote a program to calculate the great circle distance between two points on the Earth in R language by modifying the Perl version of the program written by D. Kindred of Carnegie Mellon University.

differentiated goods which are not traded on organized exchanges nor have reference prices. Rauch (1999) classified goods at the SITC 4-digit level by investigating whether each good is traded on organized exchanges or has reference prices.

In this paper I simply classified goods into two categories, i.e., differentiated goods and homogeneous goods, at the SITC 1-digit level. Goods under SITC 0-5 I classified as homogeneous goods and those under SITC 6-9 as differentiated goods. This classification seems rather simple, but it approximates Rauch's classification quite well,[10] with 72.6% of the goods under SITC 0-5 classified as Rauch's goods traded on organized exchanges or reference priced goods, and 83.1% of those under SITC 6-9 classified as his differentiated goods.

[10] Although Rauch's classification is obtainable, it cannot be used directly in the estimation in this paper because the trade data used in this paper is classified by SITC rev.3, while Rauch's classification is based on SITC rev.2.

References

Anderson J E(1979), ``A Theoretical Foundation for the Gravity Equation", *American Economic Review*, 69(1), 106-16.

Anderson J and van Wincoop, E(2002) ``Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*,March, 93(1), 170-92.

Brainard S. L(1993) ``An Empirical Assessment of the Proximity-Concentration Trade off Between Multinational Sales and Trade", *NBER Working Paper*,4580.

Deadroff A V(1998), ``Determinants of Bilateral Trade: Does Gravity Works in a Neoclassical World?", 7-28, in Frankel J. A., ed., *The regionalization of the World Economy*, University of Chicago.

Frankel J E, E Stein and S J Wei(1993) ``Continental Trading Blocks: Are They Natural, or Super-Natural?", *NBER Working Paper*,4588.

Gould, D. M.(1994) ``Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows", *Review of Economics and Statistics*. May, 76, 302-316.

Green W. H.(2000), *Econometric Analysis: Fourth edition*, Prentice Hall.

Harrigan J(2001)``Specialization and The Volume of Trade: Do The Data Obey The Laws?" , *NBER Working Paper*, 8675.

Rauch J. E(1996) ``Trade and Search: Social Capital, Sogo Shosha and Spillovers", *NBER Working Paper*, 5618.

Rauch J. E(1999) ``Networks versus Markets in International Trade" *Journal of international Economics*, 48, 7-35.

Rauch J.E and Casella A(2003) ``Overcoming Informational Barriers to International Resource allocation: Prices and Ties", *The Economic journal*,January 113, 21-42.

Rauch J.E and Trindade, V(2002) ``Ethnic Chinese Networks in International Trade", *The Review of Economics and Statistics*,February, 84(1), 116-130.

Trefler D(1995) ``The Case of the Missing Trade and Other Mysteries", *The American Economic Review*, December, 85(5), 1029-1046.