The Trade Facilitation Impact of the Chinese Diaspora

By Inmaculada Martinez-Zarzoso[1] & Robert Rudolf[2,*]

[1]Department of Economics and Center for Statistics, Georg-August Universitaet Goettingen
Platz der Goettinger Sieben 3, 37073 Goettingen, Germany and Institute of International Economics, University Jaume I, Spain

[2]Division of International Studies, Korea University
145, Anam-ro, Seongbuk-gu, Seoul 136-701, South Korea

*Corresponding author: Tel. +82-2-3290-2413, rrudolf@korea.ac.kr

This is the accepted version of the following article:

Martinez-Zarzoso, I., & Rudolf, R. (2020). The Trade Facilitation Impact of the Chinese Diaspora. *The World Economy*, **https://doi.org/10.1111/twec.12950**.

Abstract:

Using an enhanced dataset on the population share of overseas Chinese in 1970 and 1990, this paper analyzes the impact of the Chinese diaspora on facilitating China's bilateral trade during the period 1973-2013. Our findings suggest substantial trade-creation effects resulting from the presence of ethnic Chinese in the trade partner population. Diaspora impacts on Chinese bilateral imports are in general higher than those found for exports. Coethnic networks play a larger role as long as the partner country does not have an RTA with China in place. Among export sectors, effects found were strongest for food, as well as for machinery & transport equipment. In regards to imports, coethnic networks matter mostly for raw materials, machinery & transport equipment, and chemicals.

Keywords: Migration, Trade, Coethnic networks, Trade facilitation, Sectoral trade, Overseas Chinese.

JEL codes: F14, F22, O15.

Acknowledgments: We would like to thank Stefano Schiavo for valuable comments and suggestions received at the XVIII Conference in International Economics hold in La Rábida. Further thanks goes to the participants of the 2017 Singapore Economic Review Conference.

I. Introduction

With approximately US$ 2.3 trillion worth in exports of goods and services in 2015, the People's Republic of China (PRC) is by far the number one exporter in the world. In addition, China's merchandise imports stood at US$ 1.7 trillion making it the world's second largest importer closely following the United States. Understanding the determinants of Chinese bilateral trade flows thus is of vital importance given the prominent role that China plays in world trade today (Bussiere and Schnatz, 2009; Yang and Martinez-Zarzoso, 2014; Caporale et al., 2015; Johnston et al., 2015). At the same time, ethnic Chinese play an enormous role in global migration. Being the world's most populous nation and having witnessed large outmigration streams, in both past and present, gives mainland Chinese state and business actors access to a unique coethnic network spread around the world. The *overseas Chinese* with an estimated size of approximately 65 million are considered one of the largest diasporas in the world (Poston and Wong, 2016).

Rising levels of migration around the world today have sparked growing public and academic interest in the social and economic impacts of migrants. Diasporas often function as an important economic link between their source and their host countries. At least three channels have been suggested through which the presence of migrants can promote trade between source and host countries (Felbermayr et al., 2015). First, migrant networks alleviate incomplete information. They can help overcoming informal trade barriers related to language, culture, and institutions. Coethnic networks often share valuable market information, and thus help in identifying business opportunities and creating business partnerships. Second, migrant networks reduce frictions related to asymmetric information. For instance, coethnicity can raise contract enforceability since members of the same ethnic network are less likely to cheat each other. These two mechanisms constitute the trade cost channel. Third, via the preference channel, migrants boost imports to the host country if they derive higher utility from the consumption of goods made in the country of their ethnic origin (Aleksynska and Peri, 2014; Felbermayr and Toubal, 2012; Gould, 1994; Greif, 1993; Head and Ries, 1998; Munshi, 2003; Parsons and Vézina, 2018; Rauch, 2001).

Past studies that have focused on Chinese coethnic networks have established a robust effect of overseas Chinese on the facilitation of international trade (Anderson and van Wincoop, 2004; Felbermayr et al., 2010; Rauch, 2001; Rauch and Trindade, 2002) and foreign direct investment

(Gao, 2003; Gao et al., 2013; Tong, 2005).[1] Rauch (2001) and Rauch and Trindade (R&T) (2002) were the first to study the impact of overseas Chinese on bilateral trade flows. Using a gravity model for a sample of 60 countries, they were able to show that the *product* of the ethnic Chinese population shares of each two countries was positively related to these countries' bilateral trade flows in 1980 and 1990. R&T further showed that effects were stronger for differentiated than for homogeneous products, providing evidence for the hypothesis that part of the effect runs through information sharing.[2] The authors attribute their findings to Chinese coethnicity helping to lower trade costs by overcoming information barriers on the one hand and raising contract enforceability on the other. As noted by Combes et al. (2005), ethnic networks do affect bilateral trade through yet another channel, preferences for home country goods. Since R&T's landmark study, a growing number of economists have been engaged in the study of coethnic networks in international trade. Anderson and van Wincoop (2004) estimate the ad valorem tariff equivalent of informational costs implied by R&T's findings to be approximately 6 percent, a figure higher than average applied tariff rates around the world today. Felbermayr et al. (2010) re-estimate the R&T model considering multilateral resistance terms (MRT) and confirm their main findings. They also distinguish between the direct effects – involving China as a trading partner – and the indirect effects of Chinese migrants. The former are found to be sizeable, whereas the latter almost vanish in some cases, when models properly control for MRT in a cross-sectional setting. The authors further show similar trade-creation effects for other ethnic networks and find particularly strong effects for Polish, Turkish, Mexican, and Pakistani networks.

Trade-creation effects of overseas Chinese are further in line with the sociological literature, which sees diasporas "as middlemen who are active as cosmopolitan catalysts for economic transactions between global cities […] that form the backbone of the world economy" (Felbermayr et al., 2010). Sociologists have further pointed out that Chinese business networks are often built on informal personal relations based on regional connections and kinship, sometimes referred to by the popular Chinese term *Guanxi* (Folk and Jomo, 2013; Hamilton, 1996). A recent study by Priebe and Rudolf (2015) extends the discussion of economic impacts of the Chinese diaspora to

[1] Such effects have been observed not only between the host country and the PRC, but also between host country pairs.

[2] R&T define homogeneous goods as goods for which reference prices are available. In contrast, differentiated goods are defined as goods without reference prices.

aggregate economic growth in host countries. Introducing a new, enhanced dataset on the population share of overseas Chinese covering 147 host countries in 1970, the authors find that a country's initial relative endowment with overseas Chinese is positively related to subsequent economic growth of host countries. Besides enhanced investment and general TFP effects, the authors identify greater trade openness as a major growth transmission channel.

The present article's main objective is to quantify the influence of the Chinese diaspora in explaining Chinese bilateral trade flows. Using an enhanced dataset on overseas Chinese, this study estimates one-side gravity models of Chinese exports and imports, respectively. Analyses are carried out also at the sectoral level to identify heterogeneity of diaspora effects across main product groups.

The present study contributes to the existing literature in the following ways: First, by using a new dataset on ethnic Chinese compared to earlier studies, we are able to expand the number of countries included in the analysis from *63* (R&T and related studies) to *175*. This represents a big increase in the trade flows covered. Second, to the best of our knowledge, this is the first study that focuses on the question how coethnic networks affect bilateral trade flows of a major migrant *sending* country. Earlier single-country studies on the nexus of migration and trade have all used the perspective of the migrant *receiving* country (Gould (1994) for the US; Peri and Requena-Silvente (2010) for Spain; Bratti et al. (2014) for Italy; among others). Third, in contrast to earlier studies that used total trade, we estimate trade facilitation effects separately for exports and imports, accounting for the facts that the preference channel should only affect Chinese exports and that coethnic business networks can be of different importance for the sending country's exporters vs. importers. We further study the heterogeneity of effects by trading sector. Fourth, we analyze the role of informal coethnic networks as substitutes for formal trade agreements. Lastly, regarding our methodological approach, we control for time-invariant characteristics of countries and global time trends that can influence migration and trade by applying panel data techniques. In contrast to R&T and similar to Felbermayr et al. (2010), we depart from the traditional gravity model (GM) and use a Poisson pseudo-maximum likelihood (PPML) correlated random effects estimator.[3] Our

[3] Law et al. (2013) also used a correlated random effects model to estimate the effects of New Zealand's diaspora on its exports and imports but allowed for zero trade by adopting a Heckman (1979) selection model instead of the PPML estimator suggested by Santos Silva and Tenreyro (2006).

estimation technique differs from Felbermayr et al. (2010) in that, whereas they use cross-sectional estimations (for single years: 1980, 1990, 2000) we exploit the longitudinal nature of bilateral trade data and we allow for the fact that China does not export to (import from) all countries in our dataset. In particular, we use a (PPML) correlated random effects estimator with regional fixed effects that allows us to control for the unobserved heterogeneity that is region-specific and time invariant. We also use a system-GMM estimator (Blundell and Bond, 1998) that addresses the endogeneity of the Chinese network variable. We further use panel fixed-effects estimations as a robustness check.

Our main results indicate that Chinese coethnic networks indeed positively affect China's bilateral trade flows. We find substantial trade-creation effects resulting from the presence of ethnic Chinese in the trade partner population. Diaspora impacts on Chinese imports are higher than those found for exports. Coethnic networks play a larger role as long as the partner country does not have a regional trade agreement with the PRC. Sectoral analyses suggest that, among Chinese exports, diaspora effects are strongest for the sectors food, as well as machinery & transport equipment. In regards to imports, coethnic networks matter mostly for raw materials, machinery & transport equipment, and chemicals.

The remainder of this article is organized as follows. Section II describes the data and the empirical strategy used in this study. Section III presents main results, extensions to main results and sector-specific analyses, while section IV reports robustness checks. Conclusions are outlined in Section V.

II. Data and estimation method

II.1 Overseas Chinese data

Following earlier studies that use R&T-style ethnographic data, the Chinese diaspora in this study refers to the group of people that were born in or claim ancestry to China but who reside outside the People's Republic of China, Taiwan, Hong Kong, and Macau (Poston et al., 1994; Rauch and Trindade, 2002; Priebe and Rudolf, 2015; Poston and Wong, 2016). These definitions therefore involve Chinese born in mainland China but living abroad (first-generation or foreign-born migrants) as well as those Chinese that were born outside mainland China and who continue to live outside of China; in certain countries already for multiple generations. Changes in the number of overseas Chinese over time can therefore be attributed to both, fertility levels among Chinese in the host country and out-migration from China to the host country or from one host country to another.

The present study draws on country-level data on the number of overseas Chinese for 1970 and 1990 from the World Christian Encyclopedia (WCE; Barrett et al., 1982, 2001), which provides one of the most comprehensive ethnographic datasets to date and has recently been used by Priebe and Rudolf (2015) for a related analysis. WCE is a highly detailed census of churches and religions around the world which, as a by-product, offers detailed ethnographic data for most countries. WCE's ethnographic data relies on countries' population census data and additional secondary data sources.

WCE data has been frequently used in ethnic-related studies. Montalvo and Reynal-Querol (2005a; 2005b) describe it as "one of the most detailed data for ethnolinguistic diversity". In regards to measuring the size and distribution of the Chinese diaspora during historic periods, WCE's main advantage is that it has data available for many more countries compared to OCAC data used in earlier studies on the trade effects of the Chinese diaspora. For example, the 1981 World Christian Encyclopedia provides data on overseas Chinese for 183 countries. In contrast, OCAC data from its Overseas Chinese Economy Yearbook 1983/84 provides records for only 94 countries (OCAC, 1983). In particular, OCAC reports do not indicate whether countries that are not listed have no overseas Chinese population or are missing due to other reasons (political

instability, civil war, lack of census data, etc.). The inability to distinguish missing values from zeros is the main shortcoming of OCAC data.

The WCE dataset that we use in this study provides information on the number of overseas Chinese for a total of 165 countries in 1970 and 186 countries in 1990. In the following analysis, our main variable of interest is the proportion of overseas Chinese in the total population of a host country – hereafter referred to as *sharechinese*. Table 1 shows the distribution of overseas Chinese across world regions and over time using a combination of OCAC and WCE data and our own imputations. The total size of the Chinese diaspora has been constantly growing over time and across all world regions.

[Table 1 about here]

It should be noted that our definition of migration differs from many recent trade-migration studies that use data on first-generation migrants only (the so-called "foreign-born concept", see e.g. Peri and Requena-Silvente, 2010; Ö zden et al., 2011; Felbermayr and Toubal, 2012; Bratti et al., 2014). We hypothesize that trade effects of migration go well beyond the first generation of migrants, particularly with regard to the trade cost channel. Anecdotal evidence from the Chinese diaspora suggests that migrants active in trading often take a generation to build a functioning family business in their new country of residence. Thus, this wider definition of migration adds a new perspective on the trade-migration nexus to the literature.

II.2 Trade and gravity data

Bilateral Chinese exports and imports from 1973 to 2013 are taken from UN-COMTRADE[4]. Data on countries' GDP and population are drawn from the World Development Indicators Database (World Bank, 2016). Distances between capitals, as well as trade impeding or promoting factors such as common border or being landlocked are taken from the CEPII database

[4] Data were downloaded from https://comtrade.un.org/data/ on the 13th of April 2017.

(Mayer and Zignano, 2005). RTA and WTO dummy variables are –an actualized version– from De Sousa (2012). Table 2 presents summary statistics of the above variables.[5]

[Table 2 about here]

II.3 Empirical strategy

Over the past two decades, the gravity model of trade has evolved into a sophisticated tool to analyze the broad determinants of bilateral trade flows, among them a number of policy factors such as Regional Trade Agreements (RTA), trade facilitation factors, tariffs, regulations, and others (Feenstra, 2016). The gravity model has been broadly used to investigate the role played by specific policy or geographical variables in explaining bilateral trade flows. Consistent with this approach, and in order to investigate the effect of the presence of Chinese networks on Chinese trade flows, we include the variable *sharechinese* (in the trade partner country) as a "trade facilitator".

According to the underlying theory that has been reformulated and extended by Anderson and van Wincoop (2003), our model assumes constant elasticity of substitution and product differentiation by place of origin. In addition, prices differ among locations due to symmetric bilateral trade costs[6]. The reduced form of the model is specified as

$$X_{ijt} = \frac{Y_{it}Y_{jt}}{Y_t^W}\left(\frac{t_{ijt}}{P_{it}P_{jt}}\right)^{1-\sigma} \quad (1)$$

where X_{ijt} is bilateral exports from country i to country j in year t, and Y_{it}, Y_{jt} and Y_t^W are the GDPs in the exporting country, the importing country and the world in year t, respectively. t_{ijt} denotes trade costs between the exporter and the importer in year t and P_{it} and P_{jt} are price indices that account for the so-called multilateral resistance factors and are a function of the trade cost of a

[5] Following earlier studies, Hong Kong, Taiwan and Macao have been excluded from the analysis. The data that support the findings of this study are available from the corresponding author upon reasonable request.
[6] The assumption of symmetric trade costs is not required in the empirical application.

country with respect to all countries in the world. The empirical specification in log-linear form is given by

$$ln\, X_{ijt} = ln\, Y_{it} + ln\, Y_{jt} - ln\, Y_t^W + (1-\sigma) ln\, t_{ijt} - (1-\sigma) ln\, P_{it} - (1-\sigma) ln\, P_{jt} \qquad (2)$$

where ln denotes natural logarithms.

The estimation of equation (2) is not straightforward due to the presence of trade costs and multilateral resistance terms. The trade cost function is assumed to be a linear function of a number of trade barriers, namely, the time-invariant determinants of trade flows, including distance, area, common border, landlocked dummies and the time-varying RTA and WTO variables. In the recent gravity literature, multilateral resistance terms are modeled as time-varying or time-invariant country specific dummies when a full-gravity is estimated. However, for a single exporter some specificity applies.

In our empirical application, we focus exclusively on exports from (imports to) China over time for all its trading partners for which the relevant data are available. We therefore specify a one-side gravity model to explain bilateral exports and imports, in which trade partners are indexed by *j*, and years by *t*.

Since we have a single country (China) as exporter (importer) to (from) any other country in the world, and the target variable is time invariant and country specific, we opted by including regional time-invariant dummies in the main specification a way to proxy for multilateral resistance. Region-specific fixed effects are used in order to mitigate potential biases due to time-invariant unobserved heterogeneity due to cultural factors, for example. Asian countries may have communalities that affect bilateral trade between Asian countries in a different way as trade between countries in different regions. We explain below the panel data techniques used to control also for permanent country heterogeneity. We account for global trends in trade by adding common time dummies.

After dropping the *i* subscript, substitution of the trade cost function into equation (2) and addition of regional dummies, time dummies and an idiosyncratic error term suggests estimating:

$$ln\,X_{jt} = \alpha_0 + \alpha_1\,ln\,Y_{jt} + \alpha_2\,ln\,Dist_j + \alpha_3 Landlocked_j + \alpha_4 ln\,Area_j +$$
$$\alpha_5 Border_j + \alpha_6 sharechinese_{j70(90)} + \alpha_7 RTA_{jt} + \alpha_8 WTO_{jt} + \delta_r + \gamma_t + u_{jt} \quad (3)$$

where $Dist_j$ denotes geographical distance from China to country *j*. $Landlocked_j$ takes the value of one when country *j* is landlocked, and zero otherwise. $Border_j$ takes the value of one when country *j* shares a border with China, zero otherwise. $Area_j$ is the geographical area of country *j* in squared km, RTA_{jt} takes the value of one when China and the country *j* are members of the same regional trade agreement in year *t*, zero otherwise, and *WTO* takes the value of one when both countries are members of the WTO. Most important for the sake of this study, $sharechinese_j$ denotes the population share of ethnic Chinese in country *j* in 1970 or in 1990. Moreover, γ_t denotes a set of year dummies that proxy for time-variant common factors (globalization) that affect Chinese trade flows to all its partners. Finally, δ_r denote regional fixed effects and u_{jt} is the error term that is assumed to be well-behaved.

It should be reasonable to assume that the variable $sharechinese_{j70}$ is exogenous with regard to Chinese bilateral exports and imports from 1973 to 2013. As noted in Priebe and Rudolf (2015), most Chinese outmigration took place before 1952 and then again after the Open Door policies by Deng Xiaoping in 1978. The PRC's economic isolation in combination with strict migration controls between 1952 and the end of the 1970s are reason enough to believe that overseas Chinese in 1970 were not able to foresee Chinese bilateral trade during the four decades post-1973. The use of $sharechinese_{j90}$, on the other hand, is more controversial. While potentially it allows us to estimate the relationship between overseas Chinese and Chinese bilateral trade with more up-to-date population figures and for a larger sample of countries, it could clearly violate the exogeneity assumption of our estimators. To remedy this problem, we use only the trade period 1991-2013 in our estimations with this variable, in combination with other approaches to establish robustness (System GMM).

As regards the techniques used to estimate the gravity model, the main novelties are reviewed by Head and Mayer (2014). The authors discuss the main trade theories supporting the model and estimation challenges involved for correct identification of trade effects of specific economic and political factors. In our choice of estimation techniques we have to consider that we are estimating a *one-side* gravity and that the variables in the model cannot capture all influences

on China's trade. The use of panel data techniques enables us to control for permanent unobserved country-specific heterogeneity. In particular, we specify the gravity model as,

$$ln\,X_{jt} = \alpha_j + \alpha_1\,ln\,Y_{jt} + \alpha_2 ln\,Dist_j + \alpha_3 Landlocked_j + \alpha_4 ln\,Area_j + \alpha_5 Border_j + \alpha_6 sharechinese_{j70(90)} + \alpha_7 RTA_{jt} + \alpha_8 WTO_{jt} + \delta_r + \gamma_t + u_{jt} \quad (4)$$

were α_j is an unobserved country-specific effect that represents the permanent cross-country heterogeneity. The model could be estimated using a random effects approach if α_j is assumed to be uncorrelated with the regressors. Since this assumption is difficult to maintain in practice[7], we adopt a correlated random effects approach and assume that the country-specific effects are a function of the time averages of the time-variant variables:

$$\alpha_j = \beta_0 + \beta_1 \overline{\ln Y_{j.}} + \beta_2 \overline{RTA_{J.}} + \beta_3 \overline{WTO}_{j.} + \epsilon_j \quad (5)$$

Substituting equation (5) into equation (4) we obtain,

$$ln\,X_{jt} = \beta_0 + \alpha_1\,ln\,Y_{jt} + \alpha_2 ln\,Dist_j + \alpha_3 Landlocked_j + \alpha_4 ln\,Area_j + \alpha_5 Border_j + \alpha_6 sharechinese_{j70(90)} + \alpha_7 RTA_{jt} + \alpha_8 WTO_{jt} + \beta_1 \overline{ln\,Y_{j.}} + \beta_2 \overline{RTA_{J.}} + \beta_3 \overline{WTO}_{j.} \delta_r + \gamma_t + \epsilon_j + u_{jt} \quad (6)$$

Equation (6) can be estimated using random effects. Since the above approach does not allow for zero trade and given that for the regressions with sectoral data an important percent of the observations are zeros or missing, we also estimate a PPML correlated random effects model. According to it, the dependent variable is in levels. It has also the added advantage that the model is robust to heteroscedasticity in the error term.

Finally, in order to account for the potential endogeneity of the target variable (*sharechinese*) and to allow for dynamics in the dependent variable, we estimated a system-GMM model in which the first lag of the dependent variable is added as regressor and internal instruments (further lags of the dependent variable and lags of the other time variant variables) are considered as instruments

[7] It is also rejected by the regression-based Hausman test that consists on testing for the joint significance of the coefficients of the time averages of the time variant variables in equation (6).

of the potentially endogenous variables, that is the lag dependent variable and the sharechinese variable.

III. Results

III.1 Chinese diaspora and China's bilateral trade

The main results for Chinese bilateral exports and imports are presented in Tables 3 and 4, respectively. In both tables (columns 1 to 4) we use *sharechinese* in trade partner countries in 1970 (*sharechinese70*) for the trade period 1973-2013. In Tables A1 and A2 in the Appendix we employ *sharechinese* measured in 1990 (*sharechinese90*) for the period 1991-2013 instead.

The OLS estimates applied to the traditional specification of the gravity model with time fixed effects and regional fixed effects (OLS_TFE) are shown in column (1), country random effects are added to this specification (RE_TFE) in column (2) and the correlated random effects estimates, resulting from adding to the RE_TFE model the averages of the time variant variables as regressors (CRE_TFE) are shown in column (3). Finally, the results from the correlated random effects PPML model (CRE_PPML) are shown in column (4). In further robustness tests below the *sharechinese* variable will be considered as endogenously determined in a system-GMM framework.

[Table 3 about here]

[Table 4 about here]

Sharechinese70 is positive and statistically significant in all model specifications and the estimated effects found are sizeable. For bilateral exports (Table 3) the coefficients indicate that a 1-percentage point increase in *sharechinese70* in the trade partner population is associated with an increase in Chinese exports to this trade partner ranging between 1.7 to 3.1 percent in the case of *sharechinese70* (between 1.6 and 3.9 percent in the case of *sharechinese90, see Table A1*). We regard CRE_TFE and CRE_PPML as our preferred specifications, which show the most conservative estimates. Other factors that determine Chinese bilateral exports are trade partners'

market size, access to the sea, and distance to the PRC, which is largely in line with earlier studies such as Caporale et al. (2015). The models show an overall strong fit (R^2=0.87 in model 1). Most standard control variables show expected signs and levels of statistical significance.

Table 4 presents an equivalent analysis of China's bilateral imports. Results again suggest a strong role of overseas Chinese in determining China's bilateral trade. Interestingly, coefficients on the import side (Table 4) are found to be generally larger than those on the export side (Table 3). Focusing on our preferred specifications, coefficients for *sharechinese70* range between 3.1 and 4.4 (those for *sharechinese90* between 4.6 and 5.7, see Table A2). These coefficients indicate that a 1-percentage point increase in the share of ethnic Chinese in the trade partner population is associated with an increase in Chinese imports from the partner country by approximately three to six percent. The models show a good fit (R^2=0.74 in model 1), although it is lower compared to export models in Table 3.

Coethnic networks might play a more important role in countries with weak institutional environments. In a similar fashion, to the degree that trade relations are not formalized in bilateral or multilateral treaties, informal coethnic networks might play a more important role in facilitating trade between two countries. Therefore, one might want to test whether the importance of the Chinese diaspora in trade relations diminishes after countries form a regional trade agreement with the PRC.

[Table 5 about here]

Table 5 examines this hypothesis for both Chinese exports and imports by adding an interaction term between *RTA* and *sharechinese* to the gravity model (*RTA*sh70* and *RTA*sh90*). Here the CRE_PPML model has been applied, however, results are similar for other estimators. Estimation results in Table 5 confirm the above hypothesis. When countries form an RTA with the PRC, the effect of *sharechinese* diminishes by 27 percent for exports using the full trade period and *sharechinese70* (column 1) and by 25 percent when using *sharechinese90* in column (2). For imports (columns 3 and 4), the reduction is around 16 percent. These estimates suggest that, even though the formalization of trade relations tends to lower the role of informal networks in trade, the effect of overseas Chinese continues to be sizeable and highly significant after two countries form a trade agreement.

The gravity model can further be specified in a dynamic panel framework. Tables A3 and A4 in the appendix provide further evidence from System-GMM estimations which add the lagged dependent variable to the model. GMM-style instruments are used for lagged exports (imports) in columns (1) and (2) and additionally also for *sharechinese* in columns (3) and (4). Estimation results show that the lagged dependent variable is indeed significant in most models, the effect being more sizeable for exports. In contrast to imports, Chinese exports seem to follow more predictable patterns. The significance of our main variable of interest, *sharechinese*, is also confirmed by GMM estimations. Long-run coefficients of *sharechinese* for exports are 3.6 and 4.6 (columns 3 and 4 of Table A3, when *sharechinese*70/*sharechinese90* is considered endogenous), while the respective long-run coefficients for imports are 7.8 and 8.3 (columns 3 and 4 of Table A4).[8] The interaction term *rta*sharechinese* is significant in seven out of eight models of Tables A3 and A4, confirming the earlier hypothesis.

Based on our main estimates, we have calculated trade creation effects and the equivalent ad-valorem tariff reduction for Chinese exports and imports. Using the estimates from column (4) of Tables 3 and 4 for the CRE-PPML model, for Chinese exports (Table 3) we obtain a trade creation effect of 2.63 percent [9] (2.42, using Table A1, column (4) estimates) if *sharechinese70* (*90*) moves from zero to the sample average (0.015).[10] For Chinese imports the trade creation effect is significantly higher at 4.82 percent (7.08, using Table A2, column (4) estimates). These trade creation effects are equivalent to a hypothetical tariff reduction of about 0.37 (0.34) percentage points for Chinese exports and of about 0.67 (0.98) for Chinese imports (using an elasticity of substitution of 8, as in Anderson and van Wincoop (2004))[11].

[8] The AR2 test results indicate that there is autocorrelation of second order in column 4 of Table A3. However, the other three models interpreted here pass the AR2 test. All four models pass the Hansen test.

[9] This figure has been obtained using the formula: *[Exp(coef*avsh)-1] *100*; where *coef*=coefficient of sharechinese70 in column (4) of Table 3 and *avsh*=sample mean of sharechinese70 (See summary statistics in Table 2).

[10] Sample averages of 1970 sharechinese are 0.013 for the export sample and 0.017 for the import sample. Thus, we take the average of the two figures (0.015) for the calculation of trade creation and AVT reduction effects.

[11] This figure has been obtained using the formula: 100*coef**[Exp(avsh)-1] /(1-s)*; where *coef*=coefficient of sharechinese70 in column (4) of Table 3 and *avsh*=sample mean of sharechinese (See summary statistics in Table 2) and s=8 is the elasticity of substitution.

Putting our results into perspective, the magnitudes of our estimates are consistent with those found by Felbermayr et al. (2010), to the best of our knowledge the only earlier study that estimated the trade creation effects of the Chinese diaspora for the PRC's bilateral trade. Table A5 in the appendix compares our findings with those of Felbermayr et al., who found a trade creation (TC) effect of 2.4 to 8.2 percent equivalent to a reduction in ad-valorem tariffs (AVT) by 0.34 to 1.15 percentage points. We estimate this relationship over a much larger sample of countries and distinguish between export and import creation effects. We find a TC of 2.4 to 2.6 percent for exports and 4.8 to 7.1 percent for imports, respectively. That is equivalent to an AVT reduction of 0.34 to 0.37 (0.67 to 0.98) for exports (imports).

Assuming that the preference channel is only present in host countries' imports from China (Chinese exports in our study), then more products from China could be imported as the diaspora grows and the demand for Chinese products increase, but over time overseas Chinese could start producing substitutes in the destination country and the preference motive could have a negative effect on Chinese exports that could be replaced by locally produced Chinese products. In addition, one can expect that preferences for 'Chinese' products diminish for younger generations that claim Chinese ancestry but were no longer born in China, but in the host country. The cost reduction link should also be present in host countries' imports from China since Chinese exporters will have more potential links in host countries and more knowledge of the corresponding foreign markets which often give rise to significant Chinese distribution and retail networks. However, since Chinese exports mainly go to developed countries, the cost channel could be less important on average than for Chinese imports and the preference channel is probably overcompensated by a substitution channel when we consider trade in future periods as we do in our estimations.

Our results indicate that the trade creation effect is significantly higher in magnitude for Chinese imports (host countries' exports to China) than for Chinese exports (host countries' imports from China) and hence support a strong trade facilitation mechanism in place for the former. The Chinese diaspora can be expected to be better informed about the Chinese bureaucracy apparatus, contract law enforcement issues, or even the importance of bribes and this may result to be very useful when doing business in China. Additionally, being able to trust key partners in China may prove to be very valuable, as uncertainty is reduced. The diaspora has the knowledge of how to export to China, where they have family, relatives and business networks.

III.2 Sector-specific effects

Besides its overall trade expansion, in particular since its WTO accession in 2001, China has shown extraordinary trade growth in specific sectors. Among exports, machinery & transport equipment and other manufactures were the two leading sectors in 2015. The two major import sectors were primary commodities and machinery & transport equipment. Given sector-specific developments, it appears useful to examine the trade-facilitation effect of overseas Chinese through the sectoral lens.

[Table 6 about here]

Table 6 summarizes the effects of *sharechinese* on Chinese bilateral exports and imports by sector using a number of estimation techniques ranging from a log-log model with time and regional fixed effects (first row of each sector), adding random effects (second row), log-log correlated random effects (third row) and PPML with time and regional fixed effects (fourth row). Several observations can be made. First, positive and statistically significant effects prevail in almost all sectors and independent of the estimator and time period used. Second, confirming earlier results, *sharechinese90* shows somewhat stronger effects compared to *sharechinese70*. Third, with regard to exports, results suggest that coethnic networks play a particularly important role in China's overseas sales of food, and machinery & transport equipment. The strong effect on food exports is not surprising and confirms the importance of the preference channel. In contrast, the evidence of overseas Chinese promoting the PRC's exports of machinery and transport equipment points more towards the trade cost channel. Lastly, with regard to imports, overseas Chinese facilitate Chinese imports in raw materials, machinery & transport equipment, and chemicals. Given the high dependence of the Chinese economy on raw material imports, the Chinese Diaspora appears to play a vital role in the global sourcing of primary commodities. In addition, raw materials are usually imported from developing countries, in which institutions are often weak and thus there might be a stronger role of informal business links as trade facilitators.

IV. Robustness

As a first robustness check we use OCAC data on *sharechinese* to compare our main results obtained using WCE data. As mentioned earlier, OCAC data provides information on the number of overseas Chinese for significantly fewer countries as compared to WCE data, particularly for periods before 2000, but has been the standard source for data on overseas Chinese used in past studies on the Chinese diaspora. In practice, we estimate the same gravity models as presented in Tables 3 and 4 for exports and imports and replace *sharechinese1970* (WCE) with OCAC data on *sharechinese* for single years. Results of using the OCAC *sharechinese* variable for 1963 are presented in the first row of the first and second part of Table A6 in the appendix. The population share of overseas Chinese in trade partner countries has a strong effect on both Chinese bilateral exports and imports. Effects found are higher than those found earlier in Tables 3 and 4, while at the same time the number of countries covered reduced from 155 (150) using WCE data to 82 (80) using OCAC data in export (import) regressions. Table A6 shows further results from regressions using later OCAC years and it can be observed that coefficients for *sharechinese* decrease systematically when using more recent years. Therefore, estimates using earlier OCAC years might have been biased upwards due to the sample selection correlated with time that is inherent in OCAC data.

While having substantially fewer countries available, OCAC data has the potential advantage that it is measured regularly and thus one can try to use it in longitudinal form. We merged data for the Chinese diaspora for the years 1963, 1984, 2000, 2003, 2005 and 2010, creating a panel dataset in which *sharechinese* takes the value of the previous year of data available in our sample. That is, from 1963 to 1983 we assign the value of 1963, for 1984 to 1999 the value of 1984, etc. The results of estimating the gravity model with this constructed panel are reported in Table A7 in the appendix. Three models have been estimated for comparative purposes. First, results of a correlated random effects model (CRE-TFE) are presented in columns (1) and (4) for exports and imports, respectively. Second, a panel fixed-effect (FE) model, retaining only the within-variation in columns (2) and (5) and, third, the equivalent PPML FE model in columns (3) and (6). The estimations mostly confirm our main results, showing that an increase in the Chinese diaspora increases both exports and imports. However, when using panel methods, the magnitude of the effect is higher for exports than for imports, contrary to what we obtained in Tables 3 and 4. It is worth noticing that the difference in results could be due to unobserved heterogeneity that

was left uncontrolled for in the cross-sectional regressions. However, it could also be due to smaller sample size, sample selection that is correlated with time, and the resulting unbalanced nature of the OCAC data panel. OCAC data is particularly lacking for many countries before the year 2000, thus, identification using within-variation only is likely to be biased towards changes in variables in later years.

We further tried to estimate the gravity model using fixed effects panel data estimations with WCE data for 1970 and 1990 exploiting the difference between the Chinese share in both periods, however, with a single change in the target variable, the results are instable and there is not enough variability to explain changes in exports/imports.

Another sensitivity check, as suggested by Priebe and Rudolf (2015) in their analysis of diaspora effects on economic growth in host countries, consists of estimating the model only for the sample of countries with *sharechinese*$\leq$0.05, $\leq$0.03, and $\leq$0.01. This exercise helps us to verify that results are not only driven by a few countries with high population shares of overseas Chinese such as Singapore, Malaysia, or Thailand. Estimating over the restricted sample using WCE data confirms our main results. In accordance with this exercise, we also checked for nonlinearities in the effect of *sharechinese* by splitting it into several bins and also by using a quadratic specification. Results indicate that the coefficients of different bins were not statistically different and the quadratic term was not significant; thus the effects do not appear to be nonlinear.

V. Concluding remarks

This article evaluates the role of the Chinese diaspora in explaining Chinese bilateral trade flows. In order to achieve this goal, we use a new dataset on the population share of overseas Chinese and estimate one-side gravity models of Chinese exports and imports, respectively.

The present study contributes to the existing literature in the following ways: First, by using a new dataset on ethnic Chinese compared to earlier studies, we were able to expand the number of countries included in the analysis from 63 (R&T and related studies) to 175. Second, to the best of our knowledge, this is the first study that focuses on the question how coethnic networks affect bilateral trade flows of a major migrant *sending* country. Earlier single-country studies on the

nexus of migration and trade have all used the perspective of the migrant *receiving* country. Third, in contrast to earlier studies which used total trade, we estimated trade facilitation effects separately for exports and imports, accounting for the facts that the preference channel should only affect Chinese exports and that coethnic business networks can be of different importance for the sending country's exporters vs. importers. We further studied the heterogeneity of effects by trading sector. Fourth, we analyze the role of informal coethnic networks as substitutes for formal trade agreements.

Our findings suggest substantial trade-creation effects resulting from the presence of ethnic Chinese in the trade partner population. Among export sectors, effects found were strongest for food, as well as for machinery & transport equipment. In regards to imports, the largest effects were found for raw materials, machinery & transport equipment, and chemicals. Interestingly, for food products we find a higher effect for exports than for imports, indicating that for this specific sector the preference effect could play an important role. It is also worth noting that for raw materials, machinery and chemicals the trade creation effects are in general higher for imports than for exports, supporting the importance of the trade cost channel for these sectors. Putting our results into perspective, trade creation effects found are consistent with those of earlier studies such as Felbermayr et al. (2010). We are able to distinguish between Chinese export creation and import creation effects and find a trade creation of 2.4 to 2.6% and 4.8 to 7.1% respectively, equivalent to an AVT reduction of 0.34 to 0.37 and 0.67 to 0.98 percentage points.

Diaspora impacts on Chinese imports are in general higher than those found for exports. This result supports a strong trade facilitation mechanism in place particularly for Chinese imports. The Chinese diaspora can be expected to have a better idea of the Chinese bureaucracy apparatus, contract law enforcement issues, or even the importance of bribes and being able to trust key partners in China may prove to be very valuable, as uncertainty is reduced. However, this mechanism could be less relevant for Chinese exports since the bulk of Chinese exports goes to developed countries with sound institutions, and hence the trade-cost channel is probably less important.

Relatively higher trade creation effects for imports (compared to exports) of the sending country stand in contrast to findings from earlier studies (Genc et al., 2012) who more often find

the opposite to be true. A couple of explanations can be thought of. First, in contrast to the present study, earlier studies were limited by the use of first-generation migrant data. It is likely – and in line with anecdotal evidence – that first generation migrants contribute relatively more to imports from their source country than to exports to the same, indicating that the preference channel plays a larger role for first-generation migrants. Moreover, as it takes time to build up a functioning business and to be able to contribute to exports of the host country, it could be expected that the trade cost channel takes more time to be effective. If that was the case, migrants could be expected to contribute relatively more to their host countries' net bilateral exports with their country of origin as time passes. Second, overseas Chinese are often characterized by a comparatively strong emphasis on family economic success, including long working hours, and thrifty and dynamic family businesses (Bolt, 1996; Gomez et al., 2003; Folk and Jomo, 2013). Thus, for the Chinese diaspora the trade cost channel might play a more important role than the preference channel compared to other diasporas. The fact that the diaspora trade creation effect is lower for Chinese exports does not support the existence of a strong preference effect in aggregate exports. Although we cannot separate the preference effect from the trade-cost channel, a strong preference effect should lead to a higher trade creation effect on Chinese exports in comparison to Chinese imports and we find the opposite.

References

Aleksynska, M., & Peri, G. (2014). Isolating the network effect of immigrants on trade. The World Economy, 37(3), 434-455.

Anderson, J. E., & Van Wincoop, E. (2003). Gravity with gravitas: a solution to the border puzzle. The American Economic Review, 93(1), 170-192.

Anderson, J.E. & Van Wincoop, E. (2004). Trade Costs. Journal of Economic Literature, 42: 691-751.

Barrett, D. B., Kurian, G. T., & Johnson, T. M. (1982). World Christian Encyclopedia. Oxford University Press.

Barrett, D. B., Kurian, G. T., & Johnson, T. M. (2001). World Christian Encyclopedia: A comparative survey of churches and religions in the modern world, 2 vols. Oxford University Press.

Bolt, P. J. (1996). Looking to the diaspora. The overseas Chinese and China's economic development, 1978–1994. Diaspora: A Journal of Transnational Studies, 5(3), 467–496.

Blundell, R. W. and Bond, S.R. (1998), 'Initial Conditions and Moment Restrictions in Dynamic Panel Data Models', Journal of Econometrics 87, 115-143.

Bratti, M., De Benedictis, L., & Santoni, G. (2014). On the pro-trade effects of immigrants. Review of World Economics, 150(3), 557-594.

Bussière, M., & Schnatz, B. (2009). Evaluating China's integration in world trade with a gravity model based benchmark. Open Economies Review, 20(1), 85-111.

Caporale, G. M., Sova, A., & Sova, R. (2015). Trade flows and trade specialisation: The case of China. China Economic Review, 34, 261-273.

Combes, P. P., Lafourcade, M., & Mayer, T. (2005). The trade-creating effects of business and social networks: evidence from France. Journal of international Economics, 66(1), 1-29.

De Sousa, J. (2012), "The Currency Union Effect on Trade is Decreasing over Time", Economics Letters, Vol. 117/3, pp. 917-920.

Feenstra, R. C. (2016). Advanced International Trade. Theory and evidence. 2nd edition. Princeton University Press.

Felbermayr, G. J., Jung, B., & Toubal, F. (2010). Ethnic networks, information, and international trade: Revisiting the evidence. Annals of Economics and Statistics/Annales d'Économie et de Statistique, 41-70.

Felbermayr, G. J., & Toubal, F. (2012). Revisiting the trade-migration nexus: Evidence from new OECD data. World Development, 40(5), 928-937.

Felbermayr, G. J., Grossmann, V. and Kohler, W. (2015). Migration, international trade, and capital formation: Cause or effect? Ch.8 in Chriswick, R. and Miller, W. Eds. Handbook of the Economics of International Migration, Vol 1, 913-113.

Folk, B. C., & Jomo, K. S. (2013). Ethnic business: Chinese capitalism in Southeast Asia. Routledge.

Gao, T. (2003). Ethnic Chinese networks and international investment: evidence from inward FDI in China? Journal of Asian Economics, 14: 611-29.

Gao, L., Liu, X., & Zou, H. (2013). The role of human mobility in promoting Chinese outward FDI: A neglected factor?. International Business Review, 22(2), 437-449.

Genc, M., Gheasi, M., Nijkamp, P., & Poot, J. (2012). The impact of immigration on international trade: a meta-analysis. In P. Nijkamp, J. Poot, & M. Sahin (Eds.), Migration impact assessmemt. New horizons in regional science, ch. 9 (pp. 301–337). Northhampton (MA): Edward Elgar Publishing Limited.

Gomez, E. T., Hsiao, H. H. M., & Xiao, X. (Eds.) (2003). Chinese business in Southeast Asia: Contesting cultural explanations, researching entrepreneurship. Psychology Press.

Gould, D. M. (1994). Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows. The Review of Economics and Statistics, 76: 302-16.

Greif, E. (1993). Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. American Economic Review, 83(3): 525-48.

Hamilton, G. G. (Ed.). (1996). The organization foundations of Western and Chinese commerce: A historical and comparative analysis. In: Asia business networks. Berlin and New York: Walter de Gruyter.

Head, K., & Ries, J. (1998). Immigration and trade creation: econometric evidence from Canada. Canadian Journal of Economics, 47-62.

Head, K., & Mayer, T. (2014). Gravity Equations: Workhorse, Toolkit, Cookbook. In: Gopinath, G., Helpman, E. & K. Rogoff (eds.), Handbook of International Economics, Vol. 4. North Holland.

Johnston, L.A., Morgan, S.L. and Wang, Y. (2015). The Gravity of China's African Export Promise. The World Economy, 38 (6), 913-934.

Law, D., Genc, M. and Bryant, J. (2013). Trade, Diaspora and Migration to New Zealand, The World Economy, 36 (5), 582-606.

Mayer, T. and S. Zignago (2005). Market Access in Global and Regional Trade. CEPII Working Paper 2. http://www.cepii.fr/anglaisgraph/bdd/fdi.html.

Montalvo, J. G., & Reynal-Querol, M. (2005a). Ethnic diversity and economic development. Journal of Development Economics, 76(2), 293-323.

Montalvo, J. G., & Reynal-Querol, M. (2005b). Ethnic polarization, potential conflict, and civil wars. American Economic Review, 95(3), 796-816.

Munshi, K. (2003). Networks in the modern Economy: Mexican migrants in the U.S. Labor market. Quarterly Journal of Economics, 18(2): 549-599.

OCAC (1983). Overseas Chinese Economy Yearbook 1983/84 (in Chinese). Overseas Community Affairs Council, Taipei City, Taiwan, China.

Ö zden, C., Parsons, C. R., Schiff, M., & Walmsley, T. L. (2011). Where on earth is everybody? The evolution of global bilateral migration 1960–2000. The World Bank Economic Review, 25(1), 12–56.

Parsons, C., & Vézina, P. L. (2018). Migrant networks and trade: The Vietnamese boat people as a natural experiment. The Economic Journal, 128(612), F210-F234.

Peri, G., & Requena-Silvente, F. (2010). The trade creation effect of immigrants: evidence from the remarkable case of Spain. Canadian Journal of Economics/Revue canadienne d'économique, 43(4), 1433-1459.

Poston, D. L., Jr., Mao, M. X., and Mei-Yu Yu (1994). The Global Distribution of the Overseas Chinese Around 1990. Population and Development Review, 20: 631–45.

Poston , D. L., & Wong, J. H. (2016). The Chinese diaspora: The current distribution of the overseas Chinese population. Chinese Journal of Sociology, 2(3), 348-373.

Priebe, J., and Rudolf, R. (2015). Does the Chinese Diaspora Speed Up Growth in Host Countries? World Development, 76, 249-262.

Rauch, J. E. (2001). Business and Social Networks in International Trade. Journal of Economic Literature, 39: 1177-1203.

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

Santos Silva, J. and Tenreyro, S. (2006). The log of gravity. The Review of Economics and Statistics, 88(4):641–658.

Tong, S. Y. (2005). Ethnic Networks in FDI and the Impact of Institutional Development. Review of Development Economics, 9(4): 563-80.

World Bank (2016). World Development Indicators 2016. World Bank Publications.

Yang, S. and Martínez-Zarzoso, I. (2014), "A Panel Data Analysis of Trade Creation and Trade Diversion Effects: The case of ASEAN-China Free Trade Area" China Economic Review 29,138-151

Tables

Table 1: Distribution of Overseas Chinese across world regions and over time

Variable	Year	Region				
		Asia	Americas	Europe	Africa	Oceania
No of overseas Chinese (in thousands)	1970	15,360.2	1,013.6	191.0	74.3	54.6
Average share of OC across countries (%)		3.804	0.355	0.017	0.076	0.347
No of countries		36	36	33	46	8
No of overseas Chinese (in thousands)	1990	24,296.6	3,118.2	597.0	118.0	258.3
Average share of OC across countries (%)		2.648	0.475	0.046	0.066	0.678
No of countries		44	39	39	49	10
No of overseas Chinese (in thousands)	2010	28,252.1	7,249.3	1,765.5	220.9	905.4
Average share of OC across countries (%)		2.433	0.692	0.137	0.080	1.180
No of countries		44	39	39	49	11

Note: Author's calculations based on data from OCAC, WCE and authors imputations. In addition, data on the Hong Kong, Macau and Taiwan were not included in the calculations.

Table 2. Summary statistics

China's exports

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Ln exports	4,329	18.366	3.016	7.194	26.634
Ln Y_imp	4,329	23.456	2.461	16.839	30.451
Ln pop_imp	4,329	15.576	1.975	10.941	20.948
Ln dist	4,329	9.066	0.502	6.862	9.868
Border	4,329	0.055	0.228	0	1
Ln area_imp	4,329	11.593	2.446	3.912	16.654
Landlocked_imp	4,329	0.161	0.367	0	1
Sharechinese 1970 (ethnicity)	4,329	0.013	0.072	0	0.742
Sharechinese 1990 (ethnicity)	4,329	0.012	0.064	0	0.677
Sharechinese (foreign-born only)	4,275	0.001	0.007	0	0.092
WTO	4,329	0.383	0.486	0	1
RTA	4,329	0.022	0.147	0	1

China's imports

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Ln imports	3,597	17.306	3.890	0.693	25.810
Ln Y_exp	3,597	24.141	2.196	17.791	30.451
Ln pop_exp	3,597	16.045	1.784	10.878	20.948
Ln dist	3,597	9.040	0.502	6.862	9.868
Border	3,597	0.046	0.210	0	1
Ln area_exp	3,597	12.020	2.212	3.912	16.654
Landlocked_exp	3,597	0.133	0.340	0	1
Sharechinese 1970 (ethnicity)	3,597	0.017	0.087	0	0.742
Sharechinese 1990 (ethnicity)	3,597	0.014	0.066	0	0.677
Sharechinese (foreign-born only)	3,597	0.001	0.009	0	0.127
WTO	3,597	0.399	0.490	0	1
RTA	3,597	0.023	0.148	0	1

Notes: ln denotes natural logarithm; exports are in thousands of US$. Y_imp and Y_exp denote Gross Domestic Product of exporter and importer country, respectively. Pop denotes population and area denotes the geographical area of the countries. Dist is the distance between capital cities of origin and destination countries. Border (landlocked) is a dummy variable that takes the value of 1 when the trading countries share a border (do not have an exit to the sea), and zero otherwise.

Table 3: Chinese diaspora and Chinese bilateral *exports*

	(1)	(2)	(3)	(4)
Method:	OLS_TFE	RE_TFE	CRE_TFE	CRE_PPML
Dependent var:	Ln Exports			Exports
Explanatory var:	Sharechinese 1970 (ethnicity)			
Sharechinese	2.828***	3.105***	2.426***	1.731***
	[0.737]	[0.839]	[0.795]	[0.551]
Ln GDP Importer	0.850***	0.832***	0.812***	0.995***
	[0.0547]	[0.0770]	[0.0994]	[0.135]
Ln Distance	-0.474**	-0.543***	-0.482***	-0.563***
	[0.188]	[0.208]	[0.182]	[0.176]
Common Border	0.357	0.158	0.381	0.518
	[0.393]	[0.389]	[0.396]	[0.318]
Ln Area Importer	0.0622	0.0854	0.0608	0.00940
	[0.0412]	[0.0585]	[0.0411]	[0.0592]
Landlocked Importer	-0.843***	-0.816***	-0.829***	-0.932***
	[0.221]	[0.219]	[0.217]	[0.239]
RTA	-0.105	-0.250*	-0.250	-0.159*
	[0.200]	[0.152]	[0.152]	[0.0949]
WTO	0.418*	0.0251	0.00759	0.0383
	[0.220]	[0.168]	[0.172]	[0.0630]
Europe & C. Asia	-0.599**	-0.532*	-0.582*	0.139
	[0.265]	[0.317]	[0.302]	[0.356]
LA & Caribbean	-0.449	-0.367	-0.441	0.496
	[0.274]	[0.287]	[0.288]	[0.393]
MENA	0.0486	0.0293	0.140	0.328
	[0.245]	[0.245]	[0.266]	[0.281]
North America	-0.483	-0.442	-0.390	0.515
	[0.386]	[0.454]	[0.366]	[0.491]
South Asia	-0.527	-0.502	-0.531	-0.376
	[0.445]	[0.427]	[0.428]	[0.309]
Sub-Saharan Africa	0.0405	0.00390	0.0708	0.741**
	[0.229]	[0.222]	[0.241]	[0.343]
Observations	4,329	4,329	4,329	4,329
R-squared	0.866			
Number of id		155	155	155

Notes: Robust standard errors in brackets, cluster by country-pair. *** p<0.01, ** p<0.05, * p<0.1. Average of the time variant variables and time dummies are omitted to save space. East Asia and Pacific is the default region.

Table 4: Chinese diaspora and Chinese bilateral *imports*

	(1)	(2)	(3)	(4)
	OLS_TFE	RE_TFE	CRE_TFE	CRE_PPML
Dependent var.	Ln Imports			Imports
Explanatory var.	Sharechinese 1970 (ethnicity)			
Sharechinese	5.232***	5.584***	4.433***	3.136***
	[0.663]	[0.855]	[1.288]	[0.720]
Ln GDP Exporter	1.206***	1.279***	1.344***	1.073***
	[0.0701]	[0.0976]	[0.243]	[0.192]
Ln Distance	-0.502	-0.654	-0.722*	-0.807***
	[0.333]	[0.417]	[0.390]	[0.261]
Common Border	0.456	0.503	0.604	-0.0150
	[0.665]	[0.700]	[0.681]	[0.367]
Ln Area Exporter	0.289***	0.331***	0.361***	0.137**
	[0.0614]	[0.0699]	[0.0693]	[0.0573]
Landlocked Exporter	-0.0160	-0.189	-0.254	-0.0568
	[0.286]	[0.313]	[0.313]	[0.261]
RTA	0.390	-0.276	-0.321	-0.0593
	[0.352]	[0.198]	[0.207]	[0.167]
WTO	1.079**	0.384	0.181	0.0427
	[0.478]	[0.401]	[0.432]	[0.0908]
Europe & Central Asia	-0.929**	-0.971**	-0.557	-0.508
	[0.370]	[0.479]	[0.463]	[0.526]
LA & Caribbean	-1.123**	-1.024*	-0.833	-0.125
	[0.514]	[0.618]	[0.556]	[0.699]
MENA	-0.542	-0.703	-0.196	-0.539
	[0.452]	[0.559]	[0.549]	[0.699]
North America	-1.664***	-1.992***	-1.357**	-0.605
	[0.613]	[0.724]	[0.663]	[0.454]
South Asia	-1.731**	-2.189***	-2.010**	-1.596***
	[0.682]	[0.793]	[0.790]	[0.491]
Sub-Saharan Africa	-0.704	-0.488	-0.264	-0.145
	[0.431]	[0.501]	[0.478]	[0.711]
Observations	3,597	3,597	3,597	3,597
R-squared	0.740			
Number of id		150	150	150

Notes: Robust standard errors in brackets, cluster by country-pair. *** p<0.01, ** p<0.05, * p<0.1. Average variables of the time variant variables and time dummies are omitted to save space. East Asia and Pacific is the default region.

Table 5: Interaction with RTA

	(1)	(2)	(3)	(4)
Method:	CRE_PPML	CRE_PPML	CRE_PPML	CRE_PPML
Dependent var.:	*Exports*		*Imports*	
Sharechinese70	2.231***		3.655***	
	[0.574]		[0.245]	
RTA*sh70	-0.592***		-0.592**	
	[0.162]		[0.278]	
Sharechinese90		2.037***		4.166***
		[0.882]		[0.275]
RTA*sh90		-0.514**		-0.657**
		[0.160]		[0.330]
Ln GDP	0.993***	0.983***	1.072***	1.105***
	[0.137]	[0.122]	[0.104]	[0.101]
Ln Distance	-0.564***	-0.211	-0.814***	-0.859***
	[0.176]	[0.303]	[0.104]	[0.112]
Common Border	0.514	1.506***	-0.0225	0.0860
	[0.318]	[0.464]	[0.137]	[0.122]
Ln Area	0.00933	-0.0370	0.140***	0.147***
	[0.0592]	[0.0644]	[0.0204]	[0.0203]
Landlocked	-0.932***	-0.888***	-0.0551	0.108
	[0.239]	[0.240]	[0.0858]	[0.0777]
RTA	-0.00543	0.00267	0.123	0.0968
	[0.0968]	[0.0910]	[0.171]	[0.185]
WTO	0.0273	0.00732	0.0405	0.00158
	[0.0532]	[0.0489]	[0.149]	[0.164]
Observations	4,329	3,868	3,203	2,813
Number of id	155	175	150	165

Note: Robust standard errors in brackets, cluster by country-pair. *** p<0.01, ** p<0.05, * p<0.1. Average variables of the time variant variables, regional dummies and time dummies are omitted to save space. Stata command used: xtpoisson, re.

Table 6: Sectoral analysis: Chinese diaspora and Chinese bilateral trade

	Exports				***Imports***			
FOOD	sharechinese70		sharechinese90		sharechinese70		sharechinese90	
OLS-TFE	2.679***	[0.872]	3.292***	[0.711]	3.391**	[1.559]	8.551**	[3.458]
RE_TFE	3.108***	[1.044]	3.567***	[0.826]	2.507	[2.234]	6.343***	[1.825]
CRE_TFE	3.164***	[0.820]	4.010***	[0.991]	1.961	[2.575]	3.408	[2.215]
PPML	2.650***	[0.600]	2.805***	[0.684]	0.183	[0.953]	-0.0509	[0.970]
RAW MATER.	sharechinese70		sharechinese90		sharechinese70		sharechinese90	
OLS-TFE	1.594*	[0.890]	2.296**	[0.893]	6.054***	[0.897]	3.558	[2.397]
RE_TFE	2.142**	[0.976]	2.881***	[0.941]	5.665***	[1.462]	7.202***	[1.420]
CRE_TFE	1.287	[0.797]	2.149**	[1.060]	5.039***	[1.684]	5.335***	[1.548]
PPML	0.0930	[1.192]	-0.173	[1.274]	3.137***	[0.772]	3.368***	[0.908]
MACHINERY	sharechinese70		sharechinese90		sharechinese70		sharechinese90	
OLS-TFE	3.806***	[0.816]	4.339***	[0.547]	3.702***	[1.002]	-0.826	[3.307]
RE_TFE	4.049***	[0.898]	5.006***	[0.623]	4.588***	[1.603]	6.768***	[1.545]
CRE_TFE	2.884***	[0.884]	3.557***	[0.621]	2.757*	[1.456]	4.549***	[1.508]
PPML	2.295***	[0.839]	2.460***	[0.881]	5.495***	[1.487]	5.993***	[1.694]
CHEMICALS	sharechinese70		sharechinese90		sharechinese70		sharechinese90	
OLS-TFE	2.027**	[0.892]	2.687***	[0.603]	4.513***	[0.889]	2.764	[3.050]
RE_TFE	2.775***	[1.070]	4.029***	[0.704]	5.187***	[1.978]	7.782***	[1.534]
CRE_TFE	2.038***	[0.678]	2.754***	[0.568]	2.968	[1.816]	4.946***	[1.473]
PPML	1.248**	[0.586]	1.326**	[0.649]	3.813***	[0.911]	4.143***	[0.997]
OTHER MANU.	sharechinese70		sharechinese90		sharechinese70		sharechinese90	
OLS-TFE	2.347***	[0.527]	2.783***	[0.433]	3.175***	[0.785]	1.201	[2.508]
RE_TFE	1.725***	[0.591]	2.589***	[0.570]	4.320***	[1.488]	6.727***	[1.188]
CRE_TFE	2.703***	[0.694]	2.742***	[0.655]	0.395	[1.835]	2.031	[1.464]
PPML	1.322**	[0.516]	1.584***	[0.535]	1.004	[1.561]	1.469	[1.687]

Notes: Robust standard errors in brackets, clustered by country-pair. *** p<0.01, ** p<0.05, * p<0.1. All models are estimated with regional fixed effects. Regressions with sharechinese90 are for the period 1991-2013. PPML in this table is estimated with regional fixed effects, not including averages of the time-variant variables.

Appendix

Table A1: Chinese diaspora in 1990 and Chinese bilateral *exports*

	(1)	(2)	(3)	(4)
Method:	OLS_TFE	RE_TFE	CRE_TFE	CRE_PPML
Dependent var:	Ln Exports			Exports
Explanatory var:	Sharechinese 1990 (ethnicity)			
Sharechinese	3.112***	3.856***	2.740***	1.595*
	[0.506]	[0.589]	[0.555]	[0.884]
Ln GDP Importer	0.881***	0.810***	0.754***	0.985***
	[0.0500]	[0.0802]	[0.111]	[0.121]
Ln Distance	-0.493***	-0.574***	-0.478***	-0.209
	[0.177]	[0.184]	[0.184]	[0.303]
Common Border	1.042***	0.745*	1.117***	1.508***
	[0.399]	[0.436]	[0.415]	[0.464]
Ln Area Importer	0.0642	0.118**	0.0526	-0.0370
	[0.0414]	[0.0575]	[0.0426]	[0.0643]
Landlocked Importer	-0.734***	-0.793***	-0.722***	-0.889***
	[0.208]	[0.211]	[0.203]	[0.240]
RTA	-0.268*	-0.370***	-0.367***	-0.116
	[0.160]	[0.0955]	[0.0967]	[0.0830]
WTO	0.195	-0.218	-0.247	0.0156
	[0.226]	[0.186]	[0.192]	[0.0555]
Europe & C. Asia	-0.681**	-0.513**	-0.726**	-0.738
	[0.263]	[0.257]	[0.314]	[0.609]
LA & Caribbean	-0.299	-0.155	-0.342	-0.789
	[0.320]	[0.300]	[0.350]	[0.804]
MENA	-0.0327	-0.00644	-0.00697	-0.712
	[0.282]	[0.263]	[0.316]	[0.607]
North America	-0.458	-0.235	-0.410	-0.299
	[0.351]	[0.376]	[0.379]	[0.849]
South Asia	-1.003**	-0.882*	-1.063**	-1.466**
	[0.500]	[0.489]	[0.494]	[0.590]
Sub-Saharan Africa	-0.0265	-0.0837	-0.0398	-0.503
	[0.280]	[0.299]	[0.304]	[0.790]
Observations	3,868	3,868	3,868	3,868
R-squared	0.859			
Number of id		175	175	175

Notes: Robust standard errors in brackets, cluster by country-pair. *** p<0.01, ** p<0.05, * p<0.1. Average of the time variant variables and time dummies are omitted to save space. East Asia and Pacific is the default region.

Table A2: Chinese diaspora in 1990 and Chinese bilateral *imports*

	(1)	(2)	(3)	(4)
Method:	OLS_TFE	RE_TFE	CRE_TFE	CRE_PPML
Dependent var:	Ln Imports			Imports
Explanatory var:	Sharechinese 1990 (ethnicity)			
Sharechinese	6.847***	7.807***	5.697***	4.560***
	[0.768]	[1.032]	[1.142]	[1.372]
Ln GDP Importer	1.245***	1.223***	0.993***	0.668***
	[0.0669]	[0.0958]	[0.224]	[0.151]
Ln Distance	-0.960***	-1.022**	-0.921**	-0.493
	[0.345]	[0.424]	[0.364]	[0.401]
Common Border	1.114**	0.616	1.087**	0.153
	[0.476]	[0.588]	[0.543]	[0.489]
Ln Area Importer	0.338***	0.419***	0.371***	0.464***
	[0.0641]	[0.0746]	[0.0719]	[0.0641]
Landlocked Importer	-0.266	-0.396	-0.367	-0.0737
	[0.286]	[0.309]	[0.305]	[0.322]
RTA	0.0606	-0.604***	-0.610***	0.00968
	[0.387]	[0.168]	[0.159]	[0.115]
WTO	0.828**	0.0107	-0.251	0.0660
	[0.363]	[0.306]	[0.335]	[0.0805]
Europe & C. Asia	-0.456	-0.566	-0.493	-1.123
	[0.420]	[0.526]	[0.529]	[0.740]
LA & Caribbean	-0.510	-0.660	-0.677	-1.442
	[0.602]	[0.743]	[0.676]	[0.937]
MENA	0.258	-0.472	0.0127	0.144
	[0.556]	[0.729]	[0.683]	[0.777]
North America	-1.864***	-2.609***	-2.305***	-3.283***
	[0.591]	[0.837]	[0.707]	[0.920]
South Asia	-2.555***	-2.849***	-2.754***	-2.804***
	[0.548]	[0.709]	[0.671]	[0.632]
Sub-Saharan Africa	-0.209	-0.455	-0.219	-0.412
	[0.519]	[0.652]	[0.632]	[0.924]
Observations	2,813	2,813	2,813	2,813
R-squared	0.794			
Number of id		165	165	165

Notes: Robust standard errors in brackets, cluster by country-pair. *** p<0.01, ** p<0.05, * p<0.1. Average of the time variant variables and time dummies are omitted to save space. East Asia and Pacific is the default region.

Table A3: System-GMM estimations for *exports*

Dep. Variable: Ln Exports	(1) sh70_GMM	(2) sh90_GMM	(3) sh70_GMM	(4) sh90_GMM
Sharechinese long run	1.988	3.890	3.551	4.588
Sharechinese	1.732**	1.595***	2.031***	2.285***
	[0.703]	[0.563]	[0.538]	[0.529]
RTA*sharechinese	-0.695**	-0.394	-0.865***	-0.863***
	[0.320]	[0.337]	[0.261]	[0.317]
Ln Exports (t-1)	0.412***	0.590***	0.428***	0.502***
	[0.0794]	[0.0968]	[0.0774]	[0.0876]
Ln GDP Importer	0.449***	0.319***	0.435***	0.387***
	[0.0591]	[0.0801]	[0.0599]	[0.0746]
Ln Distance	-0.391***	-0.198*	-0.338***	-0.237**
	[0.117]	[0.101]	[0.116]	[0.107]
Common Border	0.0823	0.335	0.0178	0.266
	[0.234]	[0.205]	[0.267]	[0.208]
Ln Area Importer	0.0548**	0.0387	0.0468*	0.0558**
	[0.0263]	[0.0252]	[0.0254]	[0.0256]
Landlocked Importer	-0.376**	-0.278**	-0.392***	-0.396***
	[0.147]	[0.139]	[0.136]	[0.137]
RTA	0.0690	0.0330	0.104	0.113
	[0.113]	[0.0949]	[0.105]	[0.100]
WTO	0.0873	0.0378	0.174	0.100
	[0.0927]	[0.103]	[0.106]	[0.109]
Observations	4,181	3,702	4,181	3,702
Number of countries	155	175	155	175
Number of instruments	58	44	83	72
AR2 Test probability	0.109	0.00307	0.111	0.00698
Hansen probability	0.0338	0.0317	0.193	0.135

Notes: Robust standard error in brackets. *** p<0.01, ** p<0.05, * p<0.1. GMM style instruments for lagged exports in columns (1) and (2) and for lagged exports and sharechinese in columns (3) and (4). The AR2 test results indicate that there is autocorrelation of second order when sharechinese in 1990 is used in columns (2) and (4), alternative specifications using farther lags of exports did not solve the problem. Hansen test of over-identification does not pass in column (2) in Table A3 and column (1) in Table A4.

Table A4: System-GMM estimations for *imports*

Dep. Variable: Ln Imports	(1) sh70_GMM	(2) sh90_GMM	(3) sh70_GMM	(4) sh90_GMM
Sharechinese Long Run	7.059	6.993	7.762	8.332
Sharechinese	5.908***	6.993***	6.054***	7.074***
	[0.618]	[0.985]	[0.791]	[0.997]
RTA*sharechinese	-1.563**	-1.830*	-2.042**	-2.248**
	[0.717]	[0.983]	[0.917]	[0.952]
Ln Imports(t-1)	0.163**	0.125	0.220***	0.151*
	[0.0754]	[0.0950]	[0.0715]	[0.0852]
Ln GDP Exporter	0.896***	1.001***	0.836***	0.927***
	[0.0893]	[0.129]	[0.0924]	[0.115]
Ln Distance	-0.825***	-0.866***	-0.822***	-0.874***
	[0.248]	[0.220]	[0.231]	[0.217]
Common Border	-0.759*	0.0241	-0.433	-0.151
	[0.429]	[0.463]	[0.535]	[0.442]
Ln Area Exporter	0.342***	0.303***	0.325***	0.334***
	[0.0593]	[0.0638]	[0.0634]	[0.0702]
Landlocked Exporter	0.157	0.00530	0.107	-0.0880
	[0.238]	[0.262]	[0.283]	[0.272]
RTA	0.767**	0.470	0.658**	0.607*
	[0.335]	[0.346]	[0.319]	[0.321]
WTO	0.487	0.469*	0.551	0.493*
	[0.296]	[0.274]	[0.340]	[0.280]
Observations	3,151	2,516	3,151	2,516
Number of countries	147	158	147	158
Number of instruments	86	53	120	73
AR2 Test probability	0.732	0.214	0.515	0.177
Hansen probability	0.0907	0.479	0.242	0.645

Notes: Robust standard error in brackets. *** p<0.01, ** p<0.05, * p<0.1. GMM style instruments for lagged exports in columns (1) and (2) and for lagged exports and sharechinese in columns (3) and (4).

Table A5: Comparison of results to Felbermayr et al. (2010)

	Trade period	No of countries	Trade measure	Trade creation (percent)	AVT equivalent (percentage points)	Sharechinese variable	Ethnic Chinese data source
Felbermayr et al. (2010)	1980	63	Total trade	4.7	0.67	Population share of ethnic Chinese around 1980	Rauch & Trindade (2002, largely based on OCAC data)
	1990	63	Total trade	8.2	1.15	Population share of ethnic Chinese around 1990	
	2000	63	Total trade	2.4	0.34	Population share of foreign-born Chinese 2000 (first-generation migrants)	World Bank international bilateral migration stock database (Parsons et al., 2007)
Present study	1973-2013	155	Exports	2.6	0.37	Population share of ethnic Chinese around 1970	World Christian Encyclopedia (Barrett, 1982; 2001)
	1973-2013	150	Imports	4.8	0.67	Population share of ethnic Chinese around 1970	
	1991-2013	175	Exports	2.4	0.34	Population share of ethnic Chinese around 1990	World Christian Encyclopedia (Barrett, 1982; 2001)
	1991-2013	165	Imports	7.1	0.98	Population share of ethnic Chinese around 1990	

Table A6: Replications of Tables 3 and 4 with OCAC data for various years

Exports	**OLS_TFE**	**RE_TFE**	**CRE_TFE**	**CRE_PPML**	**N**	**N obs**
Sharechinese in:						
1963	6.894***	9.527***	9.405***	7.719***	82	2,356
	[0.977]	[0.812]	[1.100]	[0.943]		
1984	5.218***	6.058***	6.215***	*nc*	89	2,556
	[0.710]	[0.597]	[0.791]			
2000	3.783***	4.772***	4.707***	4.000***	109	3,511
	[0.497]	[0.512]	[0.709]	[0.516]		
2003	3.575***	4.567***	4.404***	*nc*	114	3184
	[0.465]	[0.533]	[0.703]			
2005	3.480***	4.424***	4.235***	3.739***	128	3,541
	[0.458]	[0.537]	[0.669]	[0.486]		
2010	3.323***	4.262***	4.093***	3.593***	127	3,511
	[0.431]	[0.508]	[0.642]	[0.468]		
Imports	**OLS_TFE**	**RE_TFE**	**CRE_TFE**	**CRE_PPML**	**N**	**N obs**
Sharechinese in:						
1963	8.380***	4.259***	3.695***	8.039***	80	2,254
	[1.041]	[0.791]	[1.231]	[0.652]		
1984	6.442***	2.977***	1.459	4.611***	88	2,479
	[0.752]	[0.827]	[1.310]	[0.496]		
2000	5.413***	2.733***	1.763**	4.860***	105	2,743
	[0.554]	[0.553]	[0.815]	[0.336]		
2003	5.009***	2.622***	1.725**	4.626***	109	2,825
	[0.537]	[0.535]	[0.779]	[0.320]		
2005	4.859***	2.112***	1.449*	*nc*	123	3,088
	[0.494]	[0.606]	[0.810]			
2010	4.570***	1.970***	1.304*	4.018***	122	3,074
	[0.466]	[0.571]	[0.762]	[0.264]		

Notes: Robust standard errors in brackets, cluster by country-pair. *** p<0.01, ** p<0.05, * p<0.1. Regression as in Tables 3 and 4. Only the sharechinese coefficient is displayed here. "*nc*" denotes no convergence of the PPML estimator.

Table A7: Panel data estimations with OCAC data for *exports* and *imports*

	(1)	(2)	(3)	(4)	(5)	(6)
Method:	CRE-TFE	CTFE	CTFE-PPML	CRE-TFE	CTFE	CTFE-PPML
Dependent var:	Ln Exports		Exports	Ln Imports		Imports
Explanatory var:	Sharechinese several years (ethnicity, OCAC)					
Sharechinese	5.094***	6.190**	5.163***	2.790	-2.995	2.889**
	[1.671]	[2.472]	[0.562]	[1.831]	[3.407]	[1.368]
Ln GDP Importer	0.840***	0.851***	1.040***	1.368***	1.346***	0.814***
	[0.112]	[0.113]	[0.132]	[0.233]	[0.241]	[0.154]
RTA	-0.352***	-0.353***	-0.109	-0.300	-0.296	0.0422
	[0.109]	[0.107]	[0.0683]	[0.194]	[0.194]	[0.113]
WTO	0.0322	0.0175	0.00394	0.0661	0.0142	-0.0305
	[0.234]	[0.235]	[0.0761]	[0.454]	[0.465]	[0.0819]
Ln Distance	-2.47e-05			-0.379		
	[4.03e-05]			[0.404]		
Common Border	0.478			0.714		
	[0.418]			[0.661]		
Ln Area Importer	0.0650			0.330***		
	[0.0548]			[0.0834]		
Landlocked Importer	-0.673***			-0.172		
	[0.239]			[0.353]		
Europe & C. Asia	-0.611*			-0.419		
	[0.338]			[0.468]		
LA & Caribbean	-0.802*			-1.232**		
	[0.448]			[0.594]		
MENA	-0.284			-0.401		
	[0.376]			[0.656]		
North America	-0.446			-1.552**		
	[0.420]			[0.710]		
South Asia	-0.200			-1.475**		
	[0.465]			[0.727]		
Sub-Saharan Africa	-0.242			-0.545		
	[0.370]			[0.507]		
Observations	3,002	3,002	3,002	2,676	2,676	2,676
R-squared		0.885			0.593	
Number of id	134	134	134	129	129	129

Notes: Robust standard errors in brackets, cluster by country-pair. *** p<0.01, ** p<0.05, * p<0.1. Average of the time variant variables (in column 1) and time dummies (all columns) are omitted to save space. East Asia and Pacific is the default region.