

Martínez-Zarzoso, Inmaculada; Rudolf, Robert

**Article — Published Version**

## The trade facilitation impact of the Chinese diaspora

The World Economy

**Provided in Cooperation with:**

John Wiley & Sons

*Suggested Citation:* Martínez-Zarzoso, Inmaculada; Rudolf, Robert (2020) : The trade facilitation impact of the Chinese diaspora, The World Economy, ISSN 1467-9701, Wiley, Hoboken, NJ, Vol. 43, Iss. 9, pp. 2411-2436,  
<https://doi.org/10.1111/twec.12950>

This Version is available at:

<https://hdl.handle.net/10419/230077>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>

# The trade facilitation impact of the Chinese diaspora

Inmaculada Martínez-Zarzoso<sup>1,2</sup>  | Robert Rudolf<sup>3</sup> 

<sup>1</sup>Department of Economics, University of Goettingen, Goettingen, Germany

<sup>2</sup>University Jaume I, Castellón de la Plana, Spain

<sup>3</sup>Division of International Studies, Korea University, Seoul, Korea

## Funding information

Spanish Ministry of Economy and Competitiveness, Grant/Award Number: ECO2017-83255-C3-3-P and UJI-B2017-33; Korea University Research Grant, Grant/Award Number: K1717891

## KEYWORDS

Chinese networks, correlated random-effects Poisson pseudo-maximum likelihood, gravity model, panel data, sectoral trade

## 1 | 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 (Bussière & Schnatz, 2009; Caporale, Sova, & Sova, 2015; Johnston, Morgan, & Wang, 2015; Yang & Martínez-Zarzoso, 2014). 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 & 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, Grossmann, & Kohler, 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

---

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. The *World Economy* published by John Wiley & Sons Ltd

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 & Peri, 2014; Felbermayr & Toubal, 2012; Gould, 1994; Greif, 1993; Head & Ries, 1998; Munshi, 2003; Parsons & 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 & van Wincoop, 2004; Felbermayr, Jung, & Toubal, 2010; Rauch, 2001; Rauch & Trindade, 2002) and foreign direct investment (Gao, Liu, & Zou, 2013; Gao, 2003; Tong, 2005).<sup>1</sup> 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.<sup>2</sup> 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, Lafourcade, and Mayer (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%, 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: 41). 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 & Jomo, 2013; Hamilton, 1996). A recent study by Priebe and Rudolf (2015) extends the discussion of economic impacts of the Chinese diaspora to 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 paper'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

---

<sup>1</sup>Such effects have been observed not only between the host country and the PRC, but also between host country pairs.

<sup>2</sup>R&T define homogeneous goods as goods for which reference prices are available. In contrast, differentiated goods are defined as goods without reference prices.



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 United States; Peri and Requena-Silvente (2010) for Spain; Bratti, Benedictis, and Santoni (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 versus importers. We further study the heterogeneity of effects by trading sector. Fourth, we analyse 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.<sup>3</sup> Our 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 & 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 food sector, as well as machinery and transport equipment. In regard to imports, coethnic networks matter mostly for raw materials, machinery and transport equipment, and chemicals.

The remainder of this paper is organised as follows. Section 2 describes the data and the empirical strategy used in this study. Section 3 presents main results, extensions to main results and sector-specific analyses, while section 4 reports robustness checks. Conclusions are outlined in Section 5.

## 2 | DATA AND ESTIMATION METHOD

### 2.1 | 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, Mao, & Mei-Yu, 1994; Poston

<sup>3</sup>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 Tenreiro (2006).

**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      |

*Notes:* 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.

& Wong, 2016; Priebe & Rudolf, 2015; Rauch & Trindade, 2002). 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 outmigration 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, Kurian, & Johnson, 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 rely on countries' population census data and additional secondary data sources.

World Christian Encyclopedia data have been frequently used in ethnic-related studies. Montalvo and Reynal-Querol (2005a: 298, 2005b) describe it as “one of the most detailed data for ethno-linguistic diversity.” In regard 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 provide 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.

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., Bratti



et al., 2014; Felbermayr & Toubal, 2012; Özden, Parsons, Schiff, & Walmsley, 2011; Peri & Requena-Silvente, 2010). We hypothesise 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.

## 2.2 | Trade and gravity data

Bilateral Chinese exports and imports from 1973 to 2013 are taken from UN-COMTRADE.<sup>4</sup> 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 (Mayer & Zignago, 2005). RTA and WTO dummy variables are—an actualised version—from De Sousa (2012). Table 2 presents summary statistics of the above variables.<sup>5</sup>

## 2.3 | Empirical strategy

Over the past two decades, the gravity model of trade has evolved into a sophisticated tool to analyse 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.<sup>6</sup> 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 country with respect to all countries in the world. The empirical specification in log-linear form is given by:

<sup>4</sup>Data were downloaded from <https://comtrade.un.org/data/> on 13 April, 2017.

<sup>5</sup>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.

<sup>6</sup>The assumption of symmetric trade costs is not required in the empirical application.

**TABLE 2** Summary statistics

| Variable                             | Obs   | Mean   | SD    | Min    | Max    |
|--------------------------------------|-------|--------|-------|--------|--------|
| China's exports                      |       |        |       |        |        |
| 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<br>(ethnicity)     | 4,329 | 0.013  | 0.072 | 0      | 0.742  |
| Sharechinese 1990<br>(ethnicity)     | 4,329 | 0.012  | 0.064 | 0      | 0.677  |
| Sharechinese (foreign-<br>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                      |       |        |       |        |        |
| 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<br>(ethnicity)     | 3,597 | 0.017  | 0.087 | 0      | 0.742  |
| Sharechinese 1990<br>(ethnicity)     | 3,597 | 0.014  | 0.066 | 0      | 0.677  |
| Sharechinese (foreign-<br>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. WTO (RTA) is a dummy variable that takes the value of 1 when the trading countries belong to the World Trade Organization (to the same regional trade agreement), and zero otherwise.

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

where Ln denotes natural logarithms.

The estimation of Equation (2) is not straightforward due to the presence of trade costs and multi-lateral 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 modelled 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,  $\gamma_t$  denotes a set of year dummies that proxy for time-variant common factors (globalisation) that affect Chinese trade flows to all its partners. Finally,  $\delta_r$  denotes 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 is 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:

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

where  $\alpha_j$  is an unobserved country-specific effect that represents the permanent cross-country heterogeneity. The model could be estimated using a random-effects approach if  $\alpha_j$  is assumed to be uncorrelated with the regressors. Since this assumption is difficult to maintain in practice,<sup>7</sup> 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 percentage 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 of the potentially endogenous variables, that is the lag-dependent variable and the *sharechinese* variable.

## 3 | RESULTS

### 3.1 | 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

<sup>7</sup>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).

**TABLE 3** Chinese diaspora and Chinese bilateral exports

|                         | (1)                                  | (2)                  | (3)                  | (4)                  |
|-------------------------|--------------------------------------|----------------------|----------------------|----------------------|
| <b>Method:</b>          | <b>OLS_TFE</b>                       | <b>RE_TFE</b>        | <b>CRE_TFE</b>       | <b>CRE_PPML</b>      |
| <b>Dependent var:</b>   | <b>Ln Exports</b>                    |                      |                      | <b>Exports</b>       |
| <b>Explanatory var:</b> | <b>Sharechinese 1970 (ethnicity)</b> |                      |                      |                      |
| Sharechinese            | 2.828***<br>[0.737]                  | 3.105***<br>[0.839]  | 2.426***<br>[0.795]  | 1.731***<br>[0.551]  |
| Ln GDP importer         | 0.850***<br>[0.0547]                 | 0.832***<br>[0.0770] | 0.812***<br>[0.0994] | 0.995***<br>[0.135]  |
| Ln distance             | -0.474**<br>[0.188]                  | -0.543***<br>[0.208] | -0.482***<br>[0.182] | -0.563***<br>[0.176] |
| Common border           | 0.357<br>[0.393]                     | 0.158<br>[0.389]     | 0.381<br>[0.396]     | 0.518<br>[0.318]     |
| Ln area importer        | 0.0622<br>[0.0412]                   | 0.0854<br>[0.0585]   | 0.0608<br>[0.0411]   | 0.00940<br>[0.0592]  |
| Landlocked importer     | -0.843***<br>[0.221]                 | -0.816***<br>[0.219] | -0.829***<br>[0.217] | -0.932***<br>[0.239] |
| RTA                     | -0.105<br>[0.200]                    | -0.250*<br>[0.152]   | -0.250<br>[0.152]    | -0.159*<br>[0.0949]  |
| WTO                     | 0.418*<br>[0.220]                    | 0.0251<br>[0.168]    | 0.00759<br>[0.172]   | 0.0383<br>[0.0630]   |
| Europe and C. Asia      | -0.599**<br>[0.265]                  | -0.532*<br>[0.317]   | -0.582*<br>[0.302]   | 0.139<br>[0.356]     |
| LA and Caribbean        | -0.449<br>[0.274]                    | -0.367<br>[0.287]    | -0.441<br>[0.288]    | 0.496<br>[0.393]     |
| MENA                    | 0.0486<br>[0.245]                    | 0.0293<br>[0.245]    | 0.140<br>[0.266]     | 0.328<br>[0.281]     |
| North America           | -0.483<br>[0.386]                    | -0.442<br>[0.454]    | -0.390<br>[0.366]    | 0.515<br>[0.491]     |
| South Asia              | -0.527<br>[0.445]                    | -0.502<br>[0.427]    | -0.531<br>[0.428]    | -0.376<br>[0.309]    |
| Sub-Saharan Africa      | 0.0405<br>[0.229]                    | 0.00390<br>[0.222]   | 0.0708<br>[0.241]    | 0.741**<br>[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. Average of the time-variant variables and time dummies are omitted to save space. East Asia and Pacific are the default region. \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

**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***<br>[0.663]           | 5.584***<br>[0.855]  | 4.433***<br>[1.288]  | 3.136***<br>[0.720]  |
| Ln GDP exporter         | 1.206***<br>[0.0701]          | 1.279***<br>[0.0976] | 1.344***<br>[0.243]  | 1.073***<br>[0.192]  |
| Ln distance             | -0.502<br>[0.333]             | -0.654<br>[0.417]    | -0.722*<br>[0.390]   | -0.807***<br>[0.261] |
| Common border           | 0.456<br>[0.665]              | 0.503<br>[0.700]     | 0.604<br>[0.681]     | -0.0150<br>[0.367]   |
| Ln area exporter        | 0.289***<br>[0.0614]          | 0.331***<br>[0.0699] | 0.361***<br>[0.0693] | 0.137**<br>[0.0573]  |
| Landlocked exporter     | -0.0160<br>[0.286]            | -0.189<br>[0.313]    | -0.254<br>[0.313]    | -0.0568<br>[0.261]   |
| RTA                     | 0.390<br>[0.352]              | -0.276<br>[0.198]    | -0.321<br>[0.207]    | -0.0593<br>[0.167]   |
| WTO                     | 1.079**<br>[0.478]            | 0.384<br>[0.401]     | 0.181<br>[0.432]     | 0.0427<br>[0.0908]   |
| Europe and Central Asia | -0.929**<br>[0.370]           | -0.971**<br>[0.479]  | -0.557<br>[0.463]    | -0.508<br>[0.526]    |
| LA and Caribbean        | -1.123**<br>[0.514]           | -1.024*<br>[0.618]   | -0.833<br>[0.556]    | -0.125<br>[0.699]    |
| MENA                    | -0.542<br>[0.452]             | -0.703<br>[0.559]    | -0.196<br>[0.549]    | -0.539<br>[0.699]    |
| North America           | -1.664***<br>[0.613]          | -1.992***<br>[0.724] | -1.357**<br>[0.663]  | -0.605<br>[0.454]    |
| South Asia              | -1.731**<br>[0.682]           | -2.189***<br>[0.793] | -2.010**<br>[0.790]  | -1.596***<br>[0.491] |
| Sub-Saharan Africa      | -0.704<br>[0.431]             | -0.488<br>[0.501]    | -0.264<br>[0.478]    | -0.145<br>[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. Average variables of the time-variant variables and time dummies are omitted to save space. East Asia and Pacific are the default regions. \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

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.



*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 and 3.1% in the case of *sharechinese70* (between 1.6 and 3.9% 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 ( $R^2 = .87$  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 per cent. The models show a good fit ( $R^2 = .74$  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 formalised 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 examines this hypothesis for both Chinese exports and imports by adding an interaction term between *RTA* and *sharechinese* to the gravity model ( $RTA \times sh70$  and  $RTA \times sh90$ ). 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 per cent for exports using the full trade period and *sharechinese70* (column (1)) and by 25 per cent when using *sharechinese90* in column (2). For imports (columns (3) and (4)), the reduction is around 16 per cent. These estimates suggest that, even though the formalisation 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 *sharechinese70/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).<sup>8</sup> The interaction term  $rta \times sharechinese$  is significant in seven out of eight models of Tables A3 and A4, confirming the earlier hypothesis.

<sup>8</sup>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.

TABLE 5 Interaction with RTA

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

Notes: Robust standard errors in brackets, cluster by country-pair. Average variables of the time-variant variables, regional dummies, and time dummies are omitted to save space. Stata command used: `xtpoisson, re`. \*\*\* $p < .01$ . \*\* $p < .05$ . \* $p < .1$ .

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%<sup>9</sup> (2.42, using Table A1, column (4) estimates) if *sharechinese70* (90) moves from zero to the sample average (0.015).<sup>10</sup> For Chinese imports, the trade creation effect is significantly higher at 4.82% (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

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

<sup>10</sup>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.



0.67 (0.98) for Chinese imports (using an elasticity of substitution of 8, as in Anderson and van Wincoop (2004)).<sup>11</sup>

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 2.4–8.2% equivalent to a reduction in ad-valorem tariffs (AVT) by 0.34–1.15 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 2.4–2.6% for exports and 4.8–7.1% for imports, respectively. That is equivalent to an AVT reduction of 0.34–0.37 (0.67–0.98) 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.

### 3.2 | 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 and transport equipment and other manufactures were the two leading sectors in 2015. The two major import sectors were primary commodities and machinery and transport equipment. Given sector-specific developments, it appears useful to examine the trade facilitation effect of overseas Chinese through the sectoral lens.

Table 6 summarises 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

<sup>11</sup>This figure has been obtained using the formula:  $100 \times coef \times [Exp(avsh) - 1]/(1 - s)$ , where *coef* = coefficient of *sharechinese*<sub>70</sub> 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.

**TABLE 6** Sectoral analysis: Chinese diaspora and Chinese bilateral trade

|             | Exports        |         |                |         | Imports        |         |                |         |
|-------------|----------------|---------|----------------|---------|----------------|---------|----------------|---------|
|             | sharechinese70 |         | sharechinese90 |         | sharechinese70 |         | sharechinese90 |         |
| FOOD        |                |         |                |         |                |         |                |         |
| 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.  |                |         |                |         |                |         |                |         |
| 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   |                |         |                |         |                |         |                |         |
| 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   |                |         |                |         |                |         |                |         |
| 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. |                |         |                |         |                |         |                |         |
| 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. 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. \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

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 and 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 and 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.

## 4 | 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 provide 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 have 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 regressions presented in Tables 3 and 4. 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 are 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 unstable 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.

## 5 | 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 versus importers. We further studied the heterogeneity of effects by trading sector. Fourth, we analyse 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 and transport equipment. In regard to imports, the largest effects were found for raw materials, machinery and 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 2.4–2.6% and 4.8–7.1% respectively, equivalent to an AVT reduction of 0.34–0.37 and 0.67–0.98 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, Gheasi, Nijkamp, & Poot, 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 characterised by a comparatively strong emphasis on family economic success, including long working hours, and thrifty and dynamic family businesses (Bolt, 1996; Folk & Jomo, 2013; Gomez, Hsiao, & Xiao, 2003). 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 with Chinese imports and we find the opposite.

## ACKNOWLEDGMENTS

We would like to thank the editor and two anonymous referees for their useful suggestions. We are also grateful to the participants in the XVIII International Economics Conference held in Huelva for their helpful comments.

## ORCID

Inmaculada Martínez-Zarzoso  <https://orcid.org/0000-0002-3247-8557>

Robert Rudolf  <https://orcid.org/0000-0003-1956-0457>

## 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, UK: 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* (Vol. 2). Oxford, UK: Oxford University Press.
- Blundell, R. W., & Bond, S. R. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87, 115–143.
- 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.
- 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*, 117(3), 917–920.
- Feenstra, R. C. (2016). *Advanced international trade. Theory and evidence* (2nd ed.). Princeton, NJ: Princeton University Press.

- Felbermayr, G. J., Grossmann, V., & Kohler, W. (2015). Migration, international trade, and capital formation: Cause or effect? Ch.8 In R. Chriswick, & W. Miller (Eds.), *Handbook of the economics of international migration* (Vol. 1, pp. 913–113). Amsterdam, The Netherlands: North-Holland.
- Felbermayr, G. J., Jung, B., & Toubal, F. (2010). Ethnic networks, information, and international trade: Revisiting the evidence. *Annals of Economics and Statistics*, 97/98, 41–70. <https://doi.org/10.2307/41219109>
- Felbermayr, G. J., & Toubal, F. (2012). Revisiting the trade-migration nexus: Evidence from new OECD data. *World Development*, 40(5), 928–937.
- Folk, B. C., & Jomo, K. S. (2013). *Ethnic business: Chinese capitalism in Southeast Asia*. Abingdon, UK: Routledge.
- 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.
- Gao, T. (2003). Ethnic Chinese networks and international investment: Evidence from inward FDI in China? *Journal of Asian Economics*, 14, 611–629.
- 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 assessment. New horizons in regional science*, ch. 9 (pp. 301–337). Northampton, MA: Edward Elgar Publishing Limited.
- Gomez, E. T., Hsiao, H. H. M., & Xiao, X., editors. (2003). *Chinese business in Southeast Asia: Contesting cultural explanations, researching entrepreneurship*. Hove, UK: 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, 302–316.
- Greif, E. (1993). Contract enforceability and economic institutions in early trade: The Maghribi Traders' coalition. *American Economic Review*, 83(3), 525–548.
- Hamilton, G. G. (1996). The organization foundations of Western and Chinese commerce: A historical and comparative analysis. In G. G. Hamilton (Ed.), *Asia business networks* (pp. 43–55). Berlin, Germany: Walter de Gruyter.
- Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 4, pp. 113–190). Amsterdam, The Netherlands: North Holland.
- Head, K., & Ries, J. (1998). Immigration and trade creation: Econometric evidence from Canada. *Canadian Journal of Economics*, 47–62.
- Johnston, L. A., Morgan, S. L., & Wang, Y. (2015). The gravity of China's African export promise. *The World Economy*, 38(6), 913–934.
- Law, D., Genc, M., & Bryant, J. (2013). Trade, diaspora and migration to New Zealand. *The World Economy*, 36(5), 582–606.
- Mayer, T., & Zignago, S. (2005). *Market access in global and regional trade* (CEPII Working Paper 2). Retrieved from CEPII website: <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*, 118(2), 549–599.
- OCAC (1983). *Overseas Chinese economy yearbook 1983/84 (in Chinese)*. Taipei City, Taiwan: Overseas Community Affairs Council.
- Özden, 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. R., Skeldon, R., Walmsley, T. L., & Winters, L. A. (2007). *Quantifying international migration: A database of bilateral migrant stocks*, "Policy Research Working Paper Series 4165, The World Bank. Washington DC.
- 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., & Mei-Yu, Y. U. (1994). The global distribution of the overseas Chinese around 1990. *Population and Development Review*, 20, 631–645.



- 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., & 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., & Trindade, V. (2002). Ethnic Chinese networks in international trade. *The Review of Economics and Statistics*, 84(1), 116–130.
- Santos Silva, J., & 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), 563–580.
- World Bank (2016). *World development indicators 2016*. Washington, DC: World Bank Publications.
- Yang, S., & 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.

**How to cite this article:** Martínez-Zarzoso I, Rudolf R. The trade facilitation impact of the Chinese diaspora. *World Econ.* 2020;43:2411–2436. <https://doi.org/10.1111/twec.12950>

## 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***<br>[0.506]           | 3.856***<br>[0.589]   | 2.740***<br>[0.555]   | 1.595*<br>[0.884]    |
| Ln GDP importer     | 0.881***<br>[0.0500]          | 0.810***<br>[0.0802]  | 0.754***<br>[0.111]   | 0.985***<br>[0.121]  |
| Ln distance         | -0.493***<br>[0.177]          | -0.574***<br>[0.184]  | -0.478***<br>[0.184]  | -0.209<br>[0.303]    |
| Common border       | 1.042***<br>[0.399]           | 0.745*<br>[0.436]     | 1.117***<br>[0.415]   | 1.508***<br>[0.464]  |
| Ln area importer    | 0.0642<br>[0.0414]            | 0.118**<br>[0.0575]   | 0.0526<br>[0.0426]    | -0.0370<br>[0.0643]  |
| Landlocked importer | -0.734***<br>[0.208]          | -0.793***<br>[0.211]  | -0.722***<br>[0.203]  | -0.889***<br>[0.240] |
| RTA                 | -0.268*<br>[0.160]            | -0.370***<br>[0.0955] | -0.367***<br>[0.0967] | -0.116<br>[0.0830]   |
| WTO                 | 0.195<br>[0.226]              | -0.218<br>[0.186]     | -0.247<br>[0.192]     | 0.0156<br>[0.0555]   |
| Europe and C. Asia  | -0.681**<br>[0.263]           | -0.513**<br>[0.257]   | -0.726**<br>[0.314]   | -0.738<br>[0.609]    |
| LA and Caribbean    | -0.299<br>[0.320]             | -0.155<br>[0.300]     | -0.342<br>[0.350]     | -0.789<br>[0.804]    |
| MENA                | -0.0327<br>[0.282]            | -0.00644<br>[0.263]   | -0.00697<br>[0.316]   | -0.712<br>[0.607]    |
| North America       | -0.458<br>[0.351]             | -0.235<br>[0.376]     | -0.410<br>[0.379]     | -0.299<br>[0.849]    |
| South Asia          | -1.003**<br>[0.500]           | -0.882*<br>[0.489]    | -1.063**<br>[0.494]   | -1.466**<br>[0.590]  |
| Sub-Saharan Africa  | -0.0265<br>[0.280]            | -0.0837<br>[0.299]    | -0.0398<br>[0.304]    | -0.503<br>[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. Average of the time-variant variables and time dummies are omitted to save space. East Asia and Pacific are the default regions. \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

**TABLE A2** Chinese diaspora in 1990 and Chinese bilateral imports

|                         | (1)                                  | (2)                  | (3)                  | (4)                  |
|-------------------------|--------------------------------------|----------------------|----------------------|----------------------|
| <b>Method:</b>          | <b>OLS_TFE</b>                       | <b>RE_TFE</b>        | <b>CRE_TFE</b>       | <b>CRE_PPML</b>      |
| <b>Dependent var:</b>   | <b>Ln imports</b>                    |                      |                      | <b>Imports</b>       |
| <b>Explanatory var:</b> | <b>Sharechinese 1990 (ethnicity)</b> |                      |                      |                      |
| Sharechinese            | 6.847***<br>[0.768]                  | 7.807***<br>[1.032]  | 5.697***<br>[1.142]  | 4.560***<br>[1.372]  |
| Ln GDP importer         | 1.245***<br>[0.0669]                 | 1.223***<br>[0.0958] | 0.993***<br>[0.224]  | 0.668***<br>[0.151]  |
| Ln distance             | -0.960***<br>[0.345]                 | -1.022**<br>[0.424]  | -0.921**<br>[0.364]  | -0.493<br>[0.401]    |
| Common border           | 1.114**<br>[0.476]                   | 0.616<br>[0.588]     | 1.087**<br>[0.543]   | 0.153<br>[0.489]     |
| Ln area importer        | 0.338***<br>[0.0641]                 | 0.419***<br>[0.0746] | 0.371***<br>[0.0719] | 0.464***<br>[0.0641] |
| Landlocked importer     | -0.266<br>[0.286]                    | -0.396<br>[0.309]    | -0.367<br>[0.305]    | -0.0737<br>[0.322]   |
| RTA                     | 0.0606<br>[0.387]                    | -0.604***<br>[0.168] | -0.610***<br>[0.159] | 0.00968<br>[0.115]   |
| WTO                     | 0.828**<br>[0.363]                   | 0.0107<br>[0.306]    | -0.251<br>[0.335]    | 0.0660<br>[0.0805]   |
| Europe and C. Asia      | -0.456<br>[0.420]                    | -0.566<br>[0.526]    | -0.493<br>[0.529]    | -1.123<br>[0.740]    |
| LA and Caribbean        | -0.510<br>[0.602]                    | -0.660<br>[0.743]    | -0.677<br>[0.676]    | -1.442<br>[0.937]    |
| MENA                    | 0.258<br>[0.556]                     | -0.472<br>[0.729]    | 0.0127<br>[0.683]    | 0.144<br>[0.777]     |
| North America           | -1.864***<br>[0.591]                 | -2.609***<br>[0.837] | -2.305***<br>[0.707] | -3.283***<br>[0.920] |
| South Asia              | -2.555***<br>[0.548]                 | -2.849***<br>[0.709] | -2.754***<br>[0.671] | -2.804***<br>[0.632] |
| Sub-Saharan Africa      | -0.209<br>[0.519]                    | -0.455<br>[0.652]    | -0.219<br>[0.632]    | -0.412<br>[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. Average of the time-variant variables and time dummies are omitted to save space. East Asia and Pacific are the default regions. \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

TABLE A3 System-GMM estimations for exports

|                           | (1)                  | (2)                  | (3)                  | (4)                  |
|---------------------------|----------------------|----------------------|----------------------|----------------------|
| Dep. Variable: Ln Exports | sh70_GMM             | sh90_GMM             | sh70_GMM             | sh90_GMM             |
| Sharechinese long run     | 1.988                | 3.890                | 3.551                | 4.588                |
| Sharechinese              | 1.732**<br>[0.703]   | 1.595***<br>[0.563]  | 2.031***<br>[0.538]  | 2.285***<br>[0.529]  |
| RTA × sharechinese        | -0.695**<br>[0.320]  | -0.394<br>[0.337]    | -0.865***<br>[0.261] | -0.863***<br>[0.317] |
| Ln exports ( $t - 1$ )    | 0.412***<br>[0.0794] | 0.590***<br>[0.0968] | 0.428***<br>[0.0774] | 0.502***<br>[0.0876] |
| Ln GDP importer           | 0.449***<br>[0.0591] | 0.319***<br>[0.0801] | 0.435***<br>[0.0599] | 0.387***<br>[0.0746] |
| Ln distance               | -0.391***<br>[0.117] | -0.198*<br>[0.101]   | -0.338***<br>[0.116] | -0.237**<br>[0.107]  |
| Common border             | 0.0823<br>[0.234]    | 0.335<br>[0.205]     | 0.0178<br>[0.267]    | 0.266<br>[0.208]     |
| Ln area importer          | 0.0548**<br>[0.0263] | 0.0387<br>[0.0252]   | 0.0468*<br>[0.0254]  | 0.0558**<br>[0.0256] |
| Landlocked importer       | -0.376**<br>[0.147]  | -0.278**<br>[0.139]  | -0.392***<br>[0.136] | -0.396***<br>[0.137] |
| RTA                       | 0.0690<br>[0.113]    | 0.0330<br>[0.0949]   | 0.104<br>[0.105]     | 0.113<br>[0.100]     |
| WTO                       | 0.0873<br>[0.0927]   | 0.0378<br>[0.103]    | 0.174<br>[0.106]     | 0.100<br>[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. 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 overidentification does not pass in column (2) in Table A3 and column (1) in Table A4. \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

**TABLE A4** System-GMM estimations for imports

|                                  | (1)                  | (2)                  | (3)                  | (4)                  |
|----------------------------------|----------------------|----------------------|----------------------|----------------------|
| <b>Dep. variable: Ln imports</b> | <b>sh70_GMM</b>      | <b>sh90_GMM</b>      | <b>sh70_GMM</b>      | <b>sh90_GMM</b>      |
| Sharechinese long run            | 7.059                | 6.993                | 7.762                | 8.332                |
| Sharechinese                     | 5.908***<br>[0.618]  | 6.993***<br>[0.985]  | 6.054***<br>[0.791]  | 7.074***<br>[0.997]  |
| RTA × sharechinese               | -1.563**<br>[0.717]  | -1.830*<br>[0.983]   | -2.042**<br>[0.917]  | -2.248**<br>[0.952]  |
| Ln imports ( $t - 1$ )           | 0.163**<br>[0.0754]  | 0.125<br>[0.0950]    | 0.220***<br>[0.0715] | 0.151*<br>[0.0852]   |
| Ln GDP exporter                  | 0.896***<br>[0.0893] | 1.001***<br>[0.129]  | 0.836***<br>[0.0924] | 0.927***<br>[0.115]  |
| Ln distance                      | -0.825***<br>[0.248] | -0.866***<br>[0.220] | -0.822***<br>[0.231] | -0.874***<br>[0.217] |
| Common border                    | -0.759*<br>[0.429]   | 0.0241<br>[0.463]    | -0.433<br>[0.535]    | -0.151<br>[0.442]    |
| Ln area exporter                 | 0.342***<br>[0.0593] | 0.303***<br>[0.0638] | 0.325***<br>[0.0634] | 0.334***<br>[0.0702] |
| Landlocked exporter              | 0.157<br>[0.238]     | 0.00530<br>[0.262]   | 0.107<br>[0.283]     | -0.0880<br>[0.272]   |
| RTA                              | 0.767**<br>[0.335]   | 0.470<br>[0.346]     | 0.658**<br>[0.319]   | 0.607*<br>[0.321]    |
| WTO                              | 0.487<br>[0.296]     | 0.469*<br>[0.274]    | 0.551<br>[0.340]     | 0.493*<br>[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. GMM-style instruments for lagged exports in columns (1) and (2) and for lagged exports and sharechinese in columns (3) and (4). \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .



**TABLE A5** Comparison of results to Felbermayr et al. (2010)

|                          | Trade period | No. of countries | Trade measure | Trade creation (per cent) | 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 and 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, Skeldon, Walmsley, & Winters, 2007) |
| Present study            | 1973–2013    | 155              | Exports       | 2.6                       | 0.37                               | Population share of ethnic Chinese around 1970                            | World Christian encyclopedia (Barrett et al., 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 et al., 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

|                  | OLS_TFE             | RE_TFE              | CRE_TFE             | CRE_PPML            | N   | N obs |
|------------------|---------------------|---------------------|---------------------|---------------------|-----|-------|
| Exports          |                     |                     |                     |                     |     |       |
| Sharechinese in: |                     |                     |                     |                     |     |       |
| 1963             | 6.894***<br>[0.977] | 9.527***<br>[0.812] | 9.405***<br>[1.100] | 7.719***<br>[0.943] | 82  | 2,356 |
| 1984             | 5.218***<br>[0.710] | 6.058***<br>[0.597] | 6.215***<br>[0.791] | <i>nc</i>           | 89  | 2,556 |
| 2000             | 3.783***<br>[0.497] | 4.772***<br>[0.512] | 4.707***<br>[0.709] | 4.000***<br>[0.516] | 109 | 3,511 |
| 2003             | 3.575***<br>[0.465] | 4.567***<br>[0.533] | 4.404***<br>[0.703] | <i>nc</i>           | 114 | 3,184 |
| 2005             | 3.480***<br>[0.458] | 4.424***<br>[0.537] | 4.235***<br>[0.669] | 3.739***<br>[0.486] | 128 | 3,541 |
| 2010             | 3.323***<br>[0.431] | 4.262***<br>[0.508] | 4.093***<br>[0.642] | 3.593***<br>[0.468] | 127 | 3,511 |
| Imports          |                     |                     |                     |                     |     |       |
| Sharechinese in: |                     |                     |                     |                     |     |       |
| 1963             | 8.380***<br>[1.041] | 4.259***<br>[0.791] | 3.695***<br>[1.231] | 8.039***<br>[0.652] | 80  | 2,254 |
| 1984             | 6.442***<br>[0.752] | 2.977***<br>[0.827] | 1.459<br>[1.310]    | 4.611***<br>[0.496] | 88  | 2,479 |
| 2000             | 5.413***<br>[0.554] | 2.733***<br>[0.553] | 1.763**<br>[0.815]  | 4.860***<br>[0.336] | 105 | 2,743 |
| 2003             | 5.009***<br>[0.537] | 2.622***<br>[0.535] | 1.725**<br>[0.779]  | 4.626***<br>[0.320] | 109 | 2,825 |
| 2005             | 4.859***<br>[0.494] | 2.112***<br>[0.606] | 1.449*<br>[0.810]   | <i>nc</i>           | 123 | 3,088 |
| 2010             | 4.570***<br>[0.466] | 1.970***<br>[0.571] | 1.304*<br>[0.762]   | 4.018***<br>[0.264] | 122 | 3,074 |

Notes: Robust standard errors in brackets, cluster by country-pair. Regression as in Tables 3 and 4. Only the sharechinese coefficient is displayed here. "nc" denotes no convergence of the PPML estimator. \* $p < .1$ . \*\* $p < .05$ . \*\*\* $p < .01$ .

TABLE A7 Panel data estimations with OCAC data for exports and imports

|                     | (1)  | (2)                  | (3)                 | (4)                  | (5)                 | (6)                 |
|---------------------|--|----------------------|---------------------|----------------------|---------------------|---------------------|
| Method:             | CRE-TFE                                      |                      | CTFE-PPML           | CRE-TFE              |                     | CTFE-PPML           |
| Dependent var:      | Ln Exports                                   | CTFE                 | Exports             | Ln imports           | CTFE                | Imports             |
| Explanatory var:    | Sharechinese several years (ethnicity, OCAC) |                      |                     |                      |                     |                     |
| Sharechinese        | 5.094***<br>[1.671]                          | 6.190**<br>[2.472]   | 5.163***<br>[0.562] | 2.790<br>[1.831]     | -2.995<br>[3.407]   | 2.889**<br>[1.368]  |
| Ln GDP importer     | 0.840***<br>[0.112]                          | 0.851***<br>[0.113]  | 1.040***<br>[0.132] | 1.368***<br>[0.233]  | 1.346***<br>[0.241] | 0.814***<br>[0.154] |
| RTA                 | -0.352***<br>[0.109]                         | -0.353***<br>[0.107] | -0.109<br>[0.0683]  | -0.300<br>[0.194]    | -0.296<br>[0.194]   | 0.0422<br>[0.113]   |
| WTO                 | 0.0322<br>[0.234]                            | 0.0175<br>[0.235]    | 0.00394<br>[0.0761] | 0.0661<br>[0.454]    | 0.0142<br>[0.465]   | -0.0305<br>[0.0819] |
| Ln distance         | -2.47e-05<br>[4.03e-05]                      |                      |                     | -0.379<br>[0.404]    |                     |                     |
| Common border       | 0.478<br>[0.418]                             |                      |                     | 0.714<br>[0.661]     |                     |                     |
| Ln area importer    | 0.0650<br>[0.0548]                           |                      |                     | 0.330***<br>[0.0834] |                     |                     |
| Landlocked importer | -0.673***<br>[0.239]                         |                      |                     | -0.172<br>[0.353]    |                     |                     |
| Europe and C. Asia  | -0.611*<br>[0.338]                           |                      |                     | -0.419<br>[0.468]    |                     |                     |
| LA and Caribbean    | -0.802*<br>[0.448]                           |                      |                     | -1.232**<br>[0.594]  |                     |                     |
| MENA                | -0.284<br>[0.376]                            |                      |                     | -0.401<br>[0.656]    |                     |                     |
| North America       | -0.446<br>[0.420]                            |                      |                     | -1.552**<br>[0.710]  |                     |                     |
| South Asia          | -0.200<br>[0.465]                            |                      |                     | -1.475**<br>[0.727]  |                     |                     |
| Sub-Saharan Africa  | -0.242<br>[0.370]                            |                      |                     | -0.545<br>[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. Average of the time-variant variables (in column (1)) and time dummies (all columns) are omitted to save space. East Asia and Pacific are the default regions.  $p < .1$ .  $p < .05$ .  $p < .01$ .