



AARHUS UNIVERSITY



Cover sheet

This is the accepted manuscript (post-print version) of the article.

The content in the accepted manuscript version is identical to the final published version, although typography and layout may differ.

How to cite this publication

Please cite the final published version:

Chan, J. M. L., & Zheng, H. (2022). FDI on the move: cross-border M&A and migrant networks. *Review of World Economics*, 158(3), 947-985. <https://doi.org/10.1007/s10290-021-00450-1>

Publication metadata

Title: FDI on the move: cross-border M&A and migrant networks.
Author(s): Chan, Jackie M.L. ; Zheng, Huanhuan.
Journal: Review of World Economics
DOI/Link: 10.1007/s10290-021-00450-1
Document version: Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.

FDI on the Move: Cross-border M&A and Migrant Networks

Jackie M.L. Chan* Huanhuan Zheng[†]

December 20, 2021

Abstract

Using data on Chinese outward direct investment and migrant stocks in 96 countries from 2003 to 2014, we find that migrant networks have a positive and significant impact on cross-border mergers and acquisitions (M&A), but not on greenfield investment. The migrant network effect is more pronounced for multinationals with less experience in the host country, especially for initial entrants that face greater firm, industry, and country-level information frictions. These results are robust to various estimation methods, including an instrumental variable approach that addresses potential endogeneity concerns. Our findings demonstrate the importance of knowledge spillovers from migrant networks to multinationals for facilitating entry into new markets.

JEL: F21, F22, F23

Keywords: Chinese migrant networks, cross-border mergers and acquisitions, foreign direct investment, information barriers, greenfield

*The paper has also benefited tremendously from the comments and suggestions of two anonymous referees. The authors are also grateful for the suggestions provided by Yi Lu, Han Qi, Liugang Sheng, Heiwai Tang, Beata Javorcik, and participants of the seminars at Baptist University, Lingnan University, Nanyang Technological University, the Conference on Trade and Development at the Chinese University of Hong Kong, and the Midwest Trade Conference. Jackie M.L. Chan, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210, Aarhus, Denmark, jmlchan@econ.au.dk.

[†]Huanhuan Zheng, Lee Kuan Yew School of Public Policy, National University of Singapore, 469C Bukit Timah Road, Singapore, sppzhen@nus.edu.sg.

1 Introduction

Studies have shown that social networks formed by ethnic ties and migration promote both international trade and foreign direct investment (FDI) (e.g., Rauch, 2001; Gould, 1994; Rauch and Trindade, 2002; Javorcik et al., 2011). However, the channels through which these social networks facilitate international transactions are less well understood, especially in the context of FDI. Migrants may have an information advantage or possess social capital that help foreign multinational enterprises (MNEs) overcome the barriers encountered in cross-border investment. However, such a comparative advantage may not be useful for all types of entry barriers alike, and may decay as MNEs learn from migrants and their own investment experience. Thus, this paper studies the impact of migrants in facilitating FDI and their dynamic roles in alleviating the entry barriers of multinationals.

We investigate this question using data on Chinese outward direct investment (ODI) and overseas Chinese migrant networks in 96 countries. Using regression analysis, we find that migrant networks, as measured by the stock of Chinese migrants, promote cross-border mergers and acquisitions (M&A) from China into the host country. To guard against omitted variable bias, we include a wide range of control variables. However, due to limited time variation in the migrant network size, the fixed effects model is not useful in this context to control for the unobserved heterogeneity of host countries. Instead, we employ the correlated random effects (CRE) model. By explicitly modeling the correlation between the unobserved heterogeneity and time-varying regressors, the CRE model allows for the estimation of the effects of both the time-varying and time-invariant regressors (Wooldridge, 2010). This “hybrid” model thus provides a synthesis of the fixed effects and random effects approaches.

Furthermore, we address endogeneity concerns by employing an instrumental variable (IV) approach. In the spirit of the shift-share instruments from Bartik (1991) and Burchardi et al. (2019), we employ the “leave-out push-pull” approach and construct an IV based on the interactions between the push factors from China and pull factors to the destination to rule out plausible sources of endogeneity. The IV estimation results confirm the positive impact of overseas migrants on cross-border M&A. Furthermore, we find no evidence that migrant networks are associated with greenfield FDI, an alternative entry mode in which firms set up new production facilities abroad. Both modes of FDI may attract immigrants as they expand economic activities in the host country. However, the lack of a relationship between greenfield investment and migrants suggests that the result with respect to cross-border M&A is less likely to be driven by reverse causality, in which capital flows would bring about cross-country labor movements.

To understand how migrant networks contribute to seal cross-border M&A deals, we first test their roles in transferring local knowledge to foreign MNEs. Differentiating investors’ first deal in the global markets with their subsequent investments, we find evidence for a greater impact of migrant networks on the investment of MNEs with less experience abroad. The results, which are not driven by differences in the size of investments, demonstrate the importance of migrant networks in providing information to foreign investors, and in particular, for entry into new

markets. As MNEs accumulate local knowledge and experience, the role of migrant networks diminishes for subsequent investments, though it remains statistically significant. Moreover, the migrant network effect on M&A also falls when controlling for the cumulative experience of Chinese multinationals in the host country. This suggests that migrant networks serve as a substitute source of knowledge spillovers to Chinese MNEs that have previously invested in the same destination.

Lastly, we investigate the role of migrant networks in alleviating information frictions at the firm, industry, and host-country levels. We show that the impact of migrant networks is larger for first-time entrants: (i) when either the acquirer or target company is not publicly listed, making information less transparent (e.g., financial information is not publicly disclosed); (ii) when the acquirer and target’s industries differ, so that the acquirer faces information asymmetry across industries; and (iii) in host countries with greater regulatory barriers due to business regulations and factor market frictions (e.g., labor market regulations and foreign ownership restrictions). Therefore, migrant networks are more valuable in environments with high investment barriers, whether informal or formal. Importantly, we do not find that migrants benefit inexperienced investors more when regulatory barriers are low, which again is consistent with the idea of migrants generating knowledge spillovers.

We focus on Chinese ODI in evaluating the heterogeneous roles of migrant networks for three reasons. First, Chinese ODI has grown rapidly from literally nothing to one of the world’s largest in the past two decades (see Section 2 Figure 1). Policies such as the “going out” strategy and the more recent Belt and Road Initiative have encouraged Chinese investment abroad, and this is expected to continue into the future. Moreover, with deal-level FDI data becoming available only relatively recently, the case of Chinese ODI presents an excellent opportunity to identify the investment experience of multinationals. At the start of China’s reforms in 1978, essentially all Chinese companies were operating only domestically. In contrast, the overseas expansion of firms from developed markets such as the US or Japan begins much earlier, making it difficult to track their initial entries precisely. Second, the surge in Chinese ODI reflects the common need of fast-growing emerging markets to secure long-term returns through external rebalancing, i.e., swapping low-yield international reserves for ODI. Lessons drawn from Chinese ODI practices may also be applicable to other countries with current account surpluses that accumulate foreign assets through an export-led growth strategy. Third, despite the increasing importance of capital and labor outflows from China, their comovement has largely been underexplored.¹

This paper mainly builds upon the literature on social networks and cross-border transactions. The link between ethnic or migrant networks and international trade has been demonstrated in various countries.² In particular, the results of Rauch and Trindade (2002) indicate

¹Recent works have examined other determinants of Chinese ODI, such as domestic market frictions and trade and financial linkages (e.g., Chen and Tang, 2014; Chen et al., 2019; Aizenman et al., 2018). The roles of migrant networks and different entry modes are not addressed in these studies.

²Earlier work includes Gould (1994), Head and Ries (1998), Combes et al. (2005), and Dunlevy (2006), who study the interaction of social networks and trade for US, Canada, France, and US, respectively. Some recent studies demonstrate causality by exploiting quasi-natural experiments (e.g., Steingress, 2018; Cohen et al., 2017; Parsons and Vézina, 2018). In addition, Burchardi and Hassan (2013) study the economic performance of West Germans by examining their social ties to East Germany after the fall of the Berlin Wall.

that Chinese ethnic networks provide market information on differentiated commodities, and Ottaviano et al. (2018) show that migrants can substitute for imported intermediate inputs in the UK service sectors. While studies have also found an impact of migrant networks on FDI (e.g., Buch et al., 2006; Javorcik et al., 2011), investment has generally been treated homogeneously.³ One exception is Burchardi et al. (2019), who find no difference between the investment of final goods and intermediate input producers, and along with other empirical support, also argue for the information channel of the migrant network effect. In particular, they construct and control for an index of information demand based on US county internet searches of celebrities in the migrants' origin countries, and find this variable capable of explaining FDI without a residual direct effect of migrant networks.

We contribute to this literature by demonstrating the heterogeneity of the migrant network effect on FDI along various dimensions. In particular, we rely on intuitive decompositions of the estimation sample to analyze different types of information barriers that migrant networks may alleviate. The evidence presented here is consistent with the idea that social networks transfer local knowledge to foreign entrants. This echoes with Azemar et al. (2012) that familiarity with host countries is a significant driver of FDI allocation. Importantly, our results indicate a dynamic relationship between the cross-border movements of labor and capital, in which the role of migrant networks depreciates as MNEs accumulate knowledge and experience in foreign markets. Moreover, we show a heterogeneous impact across different types of information barriers associated with information asymmetry, regulatory obstacles, and investors' entry modes. These new findings have important policy implications for both home and host countries. For example, host countries may attract FDI in the form of cross-border M&A by relaxing restrictions on immigration, and increasing the availability and transparency of information associated with foreign investment. In particular, our results suggest that new entrants of M&A will be more responsive to such measures compared to experienced investors. Positive externalities are generated in the home and host countries, with firms in the latter exposed to new technologies, management practices, and competition, leading to greater synergies and enhanced productivity.

In addition to social networks, a growing body of research finds that inter-firm linkages are also a key source of knowledge spillovers. Chaney (2014) presents a dynamic theory in which established relationships in foreign markets allow exporters to learn about new potential markets and remotely search for customers in these locations. Related work, for example, by Fernandes and Tang (2014), provides theoretical and empirical support for the influence of geographic neighbors in firms' export decisions.⁴ Like these spatial networks, the linkages of investors to social networks play a similar role in the transfer of knowledge to reduce the entry

³Other work in this area includes Kugler and Rapoport (2007) and Bhattacharya and Groznik (2008). Gao (2003) and Tong (2005) show that overseas Chinese ethnic networks have a positive correlation with Chinese inward FDI and bilateral investment, respectively. Huang et al. (2013) also analyze Chinese inward FDI, but focus on the performance of industrial firms with investment originating from ethnically Chinese economies (Hong Kong, Macau, Taiwan) versus other countries.

⁴Similarly, Kamal and Sundaram (2016) find that foreign importers are more likely to establish trade partnerships with local exporters that are in the same neighborhood of their previous partners. Bisztray et al. (2018) document evidence that firms also learn from their peers to import.

costs of cross-border transactions. We also show evidence consistent with learning dynamics in which the migrant network effect decays as MNEs accumulate knowledge and experience.

The rest of paper is organized as follows. We describe the data in Section 2 and the methodology in Section 3. Section 4 presents the empirical results, while Section 5 examines the heterogeneity of M&A to further understand the role of migrant networks in alleviating information frictions. Lastly, Section 6 concludes.

2 Data

Transaction-level data on Chinese cross-border M&A are drawn from Thomson-Reuters Security Data Company (SDC) Platinum. With comprehensive coverage of global M&A activity, this database has commonly been used for the analysis of cross-border investment (e.g., Rossi and Volpin, 2004; di Giovanni, 2005; Head and Ries, 2008; Huang and Kisgen, 2013).⁵ Our sample contains 1198 deals from China between 2003 and 2014. We choose 2003 as the initial year for two reasons. First, Chinese outward M&A only begins to grow significantly in the 21st century, with observed aggregate values exceeding 1.5 billion USD from 2003 onwards. Second, we compare our results for M&A to greenfield FDI below, for which data is available since 2003.

Figure 1(a) plots the aggregate number and observed value of cross-border M&A with Chinese MNEs as the acquiring companies. Both are clearly rising rapidly over this period of time. For example, the number of acquisitions from China is 4.65 times larger in 2014 (186 transactions) than 2003 (40). This trend is not unique to China, and is shared by other developing countries (Appendix Figure A.1(a)). In contrast, investment from developed markets exhibits no strong rising trends over these years (Appendix Figure A.1(b)).

In Figure 1(b), we classify deals based on investors' experience abroad. One group is comprised of companies' *first* M&A deal overseas, and the second group includes all *subsequent* acquisitions. At the outset of China's reforms in 1978, Chinese firms were focused on the domestic market. Because information from SDC Platinum dates back to 1981, we are able to essentially track the full history of firms' investment abroad. Thus, we aggregate first and subsequent entry separately for each acquirer. Figure 1(b) indicates that, in terms of the number of acquisitions, initial entrants are more common, as an overall 74% of M&A deals are completed by first-time investors. Moreover, the number of investment projects from companies with no experience abroad increases relatively more than that of experienced investors.

Our key explanatory variable is the overseas Chinese migrant network, and we use the stock of Chinese migrants in the host country (in logarithms) as our baseline measure (e.g., Javorcik et al., 2011). Data on bilateral migrant stocks are obtained from the United Nations (UN) Global Migration Database, available every five years from 1990 to 2015, with an additional

⁵Transaction values are not always reported due to confidentiality, so we mainly focus on the extensive margin and the counts of M&A deals. UNCTAD (2017) also maintains a database of (non-bilateral) cross-border M&A purchases at the country level. For both the number and value of Chinese M&A purchases, SDC Platinum and UNCTAD (2017) are highly correlated at 0.91 and 0.85, respectively. SDC Platinum captures 38 to 80% of the number of M&A deals annually in UNCTAD (2017), and 42 to 276% of the value.

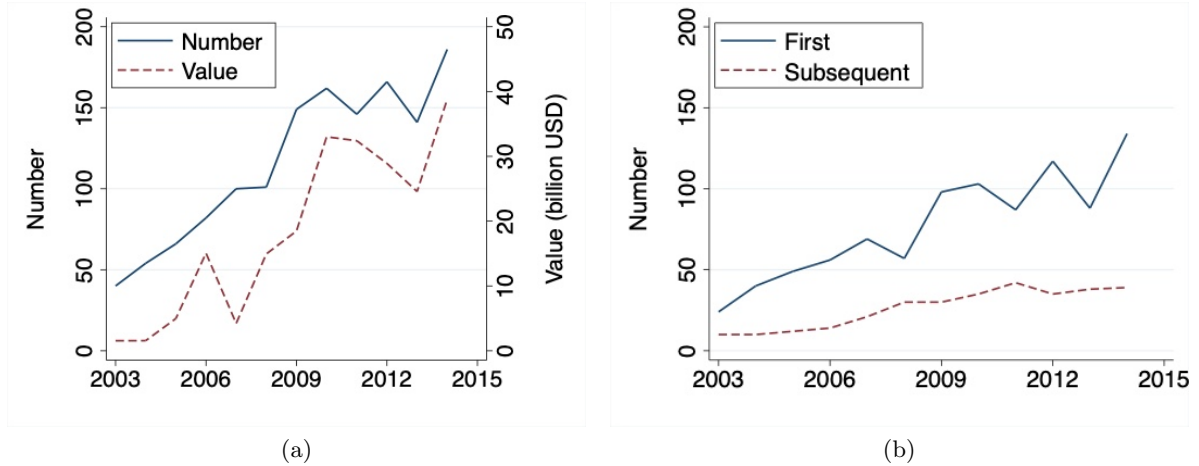


Figure 1: (a) The number and observed value of cross-border M&A deals from China; (b) the number of first and subsequent cross-border M&A deals from China.

year in 2013.⁶ For the years in which migration data are unavailable, we carry forward the value of the migrant stock until new data becomes available. For example, the Chinese migrant stock in 2006 to 2009 takes the same value as 2005. Because the growth of migrants occurs gradually, results are similar if we use linear interpolation instead. As robustness, we also consider the Chinese migrant share, i.e., the number of Chinese migrants normalized by the host country’s total population.

In our regressions, we control for a comprehensive set of host-country characteristics that may influence FDI and migration. These include traditional gravity variables like market size (i.e., GDP), income level (i.e., GDP per capita), population, geographic distance, common language, and a free trade agreement between China and the host country. Data are obtained from the World Bank World Development Indicators (WDI), CEPII, and de Sousa (2012). Also from the WDI database, we draw data on trade openness, defined as the ratio of exports plus imports to GDP (e.g., Blonigen and Piger, 2014). To capture the market’s potential, we use the growth rate of real GDP in the host country from the WDI database (e.g., Huang and Wang, 2013). The resource-seeking motives of Chinese ODI are controlled for by natural resource endowment, defined as agricultural raw materials, fuel, ores and metals exports as a share of merchandise exports (e.g., Cheung and Qian, 2009). Furthermore, we capture financial development with the private credit-to-GDP ratio (e.g., Desbordes and Wei, 2017). Lastly, following Bekaert et al. (2004), the sum of corruption, law and order, and bureaucratic quality indices from the International Country Risk Guide (ICRG) is used as a measure of institutional

⁶As stated in database documentation: “Most of the data used to estimate the international migrant stock by country or area were obtained from population censuses. Additionally, population registers and nationally representative surveys provided information on the number and composition of international migrants. In estimating the international migrant stock, international migrants have been equated with the foreign-born population whenever this information is available, which is the case in most countries or areas.” More information is available at <http://www.un.org/en/development/desa/population/migration/data/estimates2/estimates15.shtml>. Gao (2003) and Tong (2005) study the relationship between Chinese *ethnic* networks abroad and aggregate FDI using data on ethnic Chinese populations in 1990. For the overlapping year of 1990, the correlation between the population of ethnic Chinese and Chinese migrants from the UN database is high (0.68).

quality.

In our baseline estimation sample, we exclude Hong Kong, a special administrative region (SAR) of China.⁷ The statistics of this city are recorded separately from mainland China. As is well-known, there is substantial economic interaction between the two places, and importantly, the size of the Chinese migrant network in Hong Kong is also very large. Specifically, the Chinese migrant stock ranges from 1.9 to 2.3 million over our sample period, corresponding to 28 to 32% of its population. Hence, there may be concerns that Hong Kong is a potential outlier. Summary statistics for our baseline sample are provided in Appendix Table A.1.

3 Estimation strategy

Following the literature on the determinants of international trade and FDI (e.g., Bénassy-Quéré et al., 2007; Blonigen and Piger, 2014), we estimate a simple gravity model for outward cross-border M&A from China:

$$\log(MA_{it}) = \alpha + \beta \text{Migrant network}_{i,t-1} + \mathbf{X}_{i,t-1}\mathbf{\Gamma} + c_t + \varepsilon_{it}, \quad (1)$$

where MA_{it} is the number of Chinese M&A deals in host country i at year t .⁸ A positive β would be evidence in support of overseas migrant networks facilitating cross-border investment. To guard against omitted variable bias, the vector $\mathbf{X}_{i,t-1}$ includes control variables listed previously in Section 2. To mitigate the concern of reverse causality, we lag all the time-varying explanatory variables by one year (e.g., Chang, 2014). Time dummies c_t (i.e., year fixed effects) capture time-varying shocks common across destinations, and ε_{it} is the error term.

For each host country, the limited time variation of *Migrant network* precludes us from estimating its impact precisely when host-country fixed effects c_i are included in Eq. (1). While the explanatory variable is not time invariant per se, the average standard deviation of the stock of Chinese migrants across years within a country (0.31) is much smaller than the cross-sectional standard deviation (2.33). The within estimate from a fixed effect model would not be very informative on the impact of migrant networks, especially across host countries. Therefore, we instead exploit the cross-sectional variation for our econometric analysis, and treat *Migrant network* as a time-invariant variable. In order to estimate the migrant network effect, and at the same time, control for unobserved country-level heterogeneity, we employ the correlated random effects (CRE) model. Also known as the hybrid model, the CRE model unifies the fixed effects and random effects approaches (Joshi and Wooldridge, 2019). First, we rewrite Eq. (1) with host country-level unobserved heterogeneity c_i as:

$$\log(MA_{it}) = \mathbf{X}_{1,i,t-1}\mathbf{\Gamma}_1 + \mathbf{X}_{2,i}\mathbf{\Gamma}_2 + c_i + \varepsilon_{it}, \quad (2)$$

⁷The other SAR of China, Macau, and tax havens like British Virgin Islands and Cayman Islands are excluded because missing data on the control variables.

⁸Although the M&A data are available at the deal level, we are interested in the overall pattern of FDI flows and not just the size of the investment project for individual firms. Hence, we aggregate the M&A data by host country as the outcome variable. However, the detailed deal-level information allows us to decompose M&A by different characteristics and explore the channels through which migrant networks facilitate foreign investment.

where $\mathbf{X}_{1,i,t-1}$ are the time-varying, or so-called “level 1”, variables, which include all time-varying control variables and time dummies c_t . Meanwhile, $\mathbf{X}_{2,i}$ contains the time-invariant (i.e., “level 2”) variables, namely, the constant, all time-invariant control variables, and for the purposes of the estimation, *Migrant network*.

Mundlak (1978) shows that by adding the time averages of the time-varying covariates, the fixed effects estimates of $\mathbf{X}_{1,i,t-1}$ can be reproduced. In particular, this assumes that unobserved heterogeneity c_i is correlated with the time-varying observables:

$$c_i = \psi + \overline{\mathbf{X}}_{1,i}\boldsymbol{\xi} + a_i, \quad (3)$$

where $\overline{\mathbf{X}}_{1,i}$ is the time averages of $\mathbf{X}_{1,i,t-1}$, and a_i is interpreted as unobserved heterogeneity that is uncorrelated with the covariates. Substituting Eq. (3) into Eq. (2), we have

$$\log(MA_{it}) = \mathbf{X}_{1,i,t-1}\boldsymbol{\Gamma}_1 + \mathbf{X}_{2,i}\boldsymbol{\Gamma}_2 + \psi + \overline{\mathbf{X}}_{1,i}\boldsymbol{\xi} + a_i + \varepsilon_{it}. \quad (4)$$

Because a_i is assumed to be uncorrelated with the covariates, Eq. (4) can be estimated with a standard random effects model. Hence, the CRE approach is an intermediate between the fixed effects and random effects models. It allows for the estimation of time-invariant covariates $\mathbf{X}_{2,i}$ (including *Migrant network*), while at the same time controlling for unobserved heterogeneity and providing the fixed effects estimates of the time-varying covariates.

Importantly, Allison (2005), Papke and Wooldridge (2008), Wooldridge (2010), and Cameron and Trivedi (2013) show that the CRE approach works with both linear and nonlinear models. Because we have count data on the number of M&A deals, in addition to OLS, we use both a Poisson regression model and a negative binomial model. The latter is also commonly used for analyzing count data (e.g., Cameron and Trivedi, 2009; Greene, 2012), including in international economics (e.g., Blonigen and Lee, 2016), and may be potentially more efficient than the Poisson model (Cameron and Trivedi, 2013). Furthermore, Wooldridge (2019) demonstrates that the CRE approach works even when the sample is unbalanced.

4 Empirical results

4.1 Migrant networks and cross-border M&A

We begin by estimating Eq. (1) for our sample of 96 host countries. Table 1 shows a positive and statistically significant association between the size of the overseas Chinese migrant network and the number of cross-border M&A deals from China. In columns 1 and 2, we estimate our gravity model with OLS, with the former having the constant 1 added to the count of M&A transactions. Columns 3 and 4 instead fit the Poisson and negative binomial models, respectively. We have included a wide range of control variables in the regressions to mitigate omitted variable bias. In particular, we find strong correlations across the specifications between the number of deals received and the host country’s natural resource abundance and institutional quality.

Next, in Table 2, we estimate the CRE (i.e., hybrid) model and demonstrate that the results

Table 1: OLS, Poisson, and Negative Binomial Regression Estimates

Estimation method Dependent variable	OLS	OLS	Poisson	Negative binomial
	(log) MA + 1 (1)	(log) MA (2)	MA (3)	MA (4)
Migrant network	0.114*** (0.011)	0.212*** (0.056)	0.387*** (0.069)	0.334*** (0.089)
(log) GDP	0.548 (0.385)	-0.557 (1.830)	-0.564 (2.584)	-0.224 (2.184)
(log) GDP per capita	-0.537 (0.386)	0.532 (1.842)	0.874 (2.592)	0.528 (2.191)
(log) Population	-0.474 (0.387)	0.700 (1.842)	0.997 (2.605)	0.781 (2.203)
(log) Distance	0.080** (0.033)	0.168 (0.113)	0.142* (0.080)	0.052 (0.119)
Common language	0.464*** (0.131)	0.280 (0.302)	0.870 (0.557)	0.520 (0.432)
Free trade agreement	-0.182*** (0.066)	-0.202 (0.218)	-0.501 (0.308)	-0.360 (0.310)
Trade openness	0.036 (0.041)	0.070 (0.151)	0.142 (0.218)	0.384* (0.233)
Natural resources	0.215*** (0.064)	0.776 (0.465)	1.661*** (0.476)	1.675*** (0.537)
GDP growth rate	0.006 (0.005)	0.019 (0.029)	0.031 (0.024)	0.035 (0.025)
Financial development	0.158*** (0.049)	-0.016 (0.205)	-0.147 (0.247)	0.177 (0.249)
Institutional quality	0.027*** (0.009)	0.085* (0.047)	0.165*** (0.056)	0.127** (0.059)
Year FE	Y	Y	Y	Y
N	966	248	966	966
R^2	0.46	0.48		

Notes: All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country in columns 1 to 3. ***, **, * denote significance level at 1%, 5% and 10% respectively.

are generally robust even after taking into account the host countries' unobserved heterogeneity. Thus, we estimate Eq. (4) with OLS, as well as the Poisson and negative binomial models. The regressions include all the control variables, time dummies for year fixed effects, and time averages of the time-varying covariates (including time averages of the time dummies, see Wooldridge (2019)).⁹ In Panel A, we find positive and statistically significant coefficients in three out of the four columns. Although imprecisely estimated, the migrant network effect is also observed with the Poisson model. In column 4, using the negative binomial model, we find that all else equal, a 1% increase in the Chinese migrant network size raises the number of M&A deals to the host country by approximately 0.233%. In Panel B, we show that similar results are obtained when the potential outlier of Hong Kong is included in the sample. As expected, given the high

⁹Standard errors are clustered by host country for the linear and Poisson models to allow for the correlation of error terms across years within each destination. For the negative binomial model, the estimation of clustered standard errors in Stata is not currently feasible with the CRE approach (see the documentation of the `xtnbreg` command). However, as noted by Cameron and Trivedi (2009, p. 627), using the negative binomial model “may lead to improved efficiency in estimation and a default estimate of the VCE ([variance–covariance matrix of the estimates]) that should be much closer to the cluster-robust estimate of the VCE, unlike the Poisson panel commands.”

Table 2: Estimates Using the Correlated Random Effects Model

Panel A: Baseline estimates				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Migrant network	0.072*** (0.025)	0.137* (0.073)	0.176 (0.275)	0.233*** (0.070)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	966	248	966	966
Panel B: Including Hong Kong				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Migrant network	0.074** (0.028)	0.161* (0.097)	0.241 (0.199)	0.276*** (0.071)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	978	260	978	978
Panel C: Alternative measure of the migrant network				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Chinese share of migrants	0.372*** (0.099)	0.247*** (0.071)	0.331*** (0.116)	0.337*** (0.119)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	966	248	966	966

Notes: The set of control variables is described in the text. Standard errors in parentheses are clustered by host country in columns 1 to 3. ***, **, * denote significance level at 1%, 5% and 10% respectively.

correlation of bilateral economic activity between Hong Kong and mainland China as well as the large number of Chinese migrants living in Hong Kong, the estimated effects are slightly larger in magnitude. In Panel C, we employ an alternative measure of the migrant network, namely, the Chinese share of migrants (i.e., the number of Chinese migrants divided by host-country population). Across all four columns, the coefficient on our regressor of interest is positive and statistically significant.

For our baseline CRE model in Table 2 Panel A, we present the full results with the coefficients of all covariates in Appendix Table A.2. For comparison, in Appendix Table A.3, we also estimate the CRE model by treating *Migrant network* as a time-varying covariate. Thus, this particular specification includes the time-averaged variable, $\overline{Migrant\ network}$, which captures the cross-sectional variation across host countries. Under this alternative approach, the coefficient of *Migrant network* provides the within estimate, and the coefficient of $\overline{Migrant\ network}$ the between estimate. Appendix Table A.3 shows statistically insignificant coefficients on

Migrant network, but positive and statistically significant coefficients on $\overline{Migrant\ network}$ across all four estimation methods, including the Poisson model. Therefore, as expected, it is the cross-sectional variation across host countries that is driving the observed migrant network effect in our baseline estimates from Table 2. Keeping this interpretation of our results in mind, we continue treating *Migrant network* as a time-invariant variable.

4.2 Robustness

In Table 3, we conduct a series of robustness checks to confirm the positive association between migrant networks and cross-border M&A. First, in Panel A, we combine the outbound deals of China and Hong Kong, which are counted separately in SDC Platinum. Hong Kong is a platform from which many mainland Chinese companies invest abroad, so deals from this city may also benefit from the presence of the (mainland) Chinese migrants in the host country. For both the Poisson and negative binomial models, we obtain smaller coefficients compared to Table 2 Panel A. This may suggest that migrants from mainland China are slightly less helpful for investors from Hong Kong.

Next, in Panel B, we demonstrate robustness and alleviate concerns of reverse causality by subtracting recent *flows* of Chinese migrants from their stocks in the host countries. Data on the bilateral flows of migrants are obtained from Abel and Cohen (2019). Employing the same UN Global Migration Database, the authors apply six different methods to compute bilateral migration flows between 1990 and 2015. For instance, one method is stock differencing, in which the bilateral stock at one point in time is subtracted by the lagged stock. However, this simple method has difficulties with, for example, reconciling negative flows. Instead, we rely on the flow measure computed by Abel and Cohen (2019) using a demographic accounting approach in a closed system, with a pseudo-Bayesian method to estimate missing flows. Based on their validation exercises, this is one of their two preferred approaches. In particular, it improves upon methods such as stock differencing by estimating migration flows to match changes in the reported bilateral stocks of migrants, and births and deaths (Abel and Cohen, 2019). If the direction of causality is reversed, such that ODI drives migration, then the migrant stock would rise due to the influx of new migrants into the host country. Hence, we can mitigate concerns of reverse causality by subtracting recent Chinese migration flows (i.e., in the last 5 years) from the Chinese migrant stock. This alternative measure of the migrant network, denoted as $Migrant\ network^{-Flows}$, is based on its historical size. In Panel B, the migrant network effect remains robust with a comparable magnitude. Similar qualitative results are also obtained with the other five methods from Abel and Cohen (2019); these are available upon request.

In Panel C, we conduct our analysis at the industry level to show that our results are not driven by the industry composition of M&A across host countries. In the CRE model, “level 1” is now defined at the host country-industry level, so we average the covariates over time within country-industry pairs. We classify transactions into one of ten divisions in the SIC codes hierarchy. Because many more regressors must now be included in the regression, we use this rather broad classification. However, we note the negative binomial model delivers a similar

Table 3: Robustness Checks

Panel A: Aggregating China and Hong Kong FDI				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Migrant network	0.075** (0.035)	0.153** (0.070)	0.078 (0.417)	0.167*** (0.057)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	966	361	966	966
Panel B: Subtracting recent flows				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Migrant network ^{-Flows}	0.078*** (0.024)	0.219*** (0.067)	0.201 (0.193)	0.222*** (0.068)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	923	233	923	923
Panel C: Industry-level regressions				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Migrant network	0.009* (0.005)	0.060* (0.030)	0.176 (2.947)	0.169** (0.068)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	9,660	449	9,660	9,660
Panel D: Other countries' FDI				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Migrant network	0.011 (0.052)	-0.003 (0.053)	0.013 (0.053)	0.008 (0.024)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	966	940	966	966
Panel E: Taiwanese migrant network				
Estimation method	OLS	OLS	Poisson	Negative binomial
Dependent variable	(log) MA+1	(log) MA	MA	MA
	(1)	(2)	(3)	(4)
Taiwanese migrant network	0.021 (0.020)	-0.001 (0.033)	-0.016 (0.034)	-0.005 (0.031)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	966	248	966	966

Notes: The set of control variables is described in the text. Standard errors in parentheses are clustered by host country in columns 1 to 3. ***, **, * denote significance level at 1%, 5% and 10% respectively.

finding when we utilize a narrower 2-digit SIC industry classification. Panel C confirms that the positive relationship between migrant networks and cross-border M&A is also observed at the industry level.

To demonstrate that our results are not driven by the spurious correlation of migration and FDI from all origin countries, we show that Chinese migrants are not positively associated with the ODI from other countries into the host nation. Thus, in Panel D, we replace the dependent variable with non-Chinese cross-border M&A in the host country. Indeed, the effects are weak and statistically insignificant, which suggests that Chinese migrants serve as information networks for investment from China only.

Instead of examining non-Chinese FDI, in Panel E, we consider migrants from another country. Specifically, we replace the network of migrants from mainland China with the Taiwanese migrant network. This tests the idea that Taiwanese migrants may play the same role as the mainland Chinese migrants in the host country, given the two countries share a common language and cultural similarities. While Taiwan is not included in the UN Global Migration Database, it is part of the World Bank Global Bilateral Migration Database, available every 10 years from 1960 to 2000. Because we mainly exploit the cross-sectional variation for our analysis and there is generally strong correlation between the past and current migrant network size, we use data from the year 2000. In Panel E, the estimated coefficients of the Taiwanese migrant network are small in magnitude and statistically insignificant. This implies that there are barriers and obstacles that are specific to Chinese investors, and the Taiwanese migrants are not helpful in alleviating them. For example, the Chinese investors may be impeded by difficulties in access to finance or differences in regulations between the host country and China. We show evidence of such channels in Section 5.3 below.

4.3 Endogeneity and an instrumental variable approach

Thus far, we have documented a positive and significant relationship between migrant networks and M&A from China in the host country. Besides controlling for a number of host-country characteristics to mitigate omitted variable bias, we have employed the CRE model to account for host-country unobserved heterogeneity. However, one may still be concerned with potential endogeneity. To address these concerns, we take two approaches. First, we use an instrumental variable (IV) approach to formally tackle potential endogeneity issues. Second, we explore the relationship between migrants and Chinese greenfield ODI, an alternative entry mode. If our main finding is instead driven by Chinese ODI promoting migration to the host country, one would expect a positive relationship between any mode of Chinese ODI and its overseas migrant networks. Otherwise, we provide evidence inconsistent with reverse causality. Indeed, we show below that there is no impact of Chinese migrants on Chinese greenfield ODI, even when we instrument for the migrant network in our regressions.

4.3.1 Instrumental variable approach

We address concerns of endogeneity here using an IV approach. In particular, we construct an instrument for *Migrant network*, i.e., the logarithm of the Chinese migrant stock in the host country, by predicting the migrant stock from the interaction of push and pull factors. Exogenous variation is introduced by exploiting global bilateral migration flows. The computation of the *level* of the predicted migrant stock is similar in spirit to the shift-share instrument of Bartik (1991), and the ratio of the contemporaneous to lagged migrant stock may be decomposed as an inner product of two terms (Goldsmith-Pinkham et al., 2020). However, we take the logarithm of the predicted migrant stock, and therefore, the construction of our instrumental variable most closely resembles the “leave-out push-pull” approach in Burchardi et al. (2019).¹⁰

More formally, the level of the Chinese migrant stock in country i at year t , $Migrant_{i,t}^{CHN}$, can be written as the migrant stock in year $t-1$ times one plus the growth rate, or equivalently:

$$Migrant_{i,t}^{CHN} = Migrant_{i,t-1}^{CHN} \times \left(1 + \underbrace{\frac{\Delta Migrant_{WLD,t}^{CHN}}{Migrant_{WLD,t-1}^{CHN}}}_{\widetilde{Push}_{i,t}} \times \underbrace{\frac{\frac{\Delta Migrant_{i,t}^{CHN}}{Migrant_{i,t-1}^{CHN}}}{\frac{\Delta Migrant_{WLD,t}^{CHN}}{Migrant_{WLD,t-1}^{CHN}}}}_{\widetilde{Pull}_{i,t}} \right), \quad (5)$$

where $\Delta Migrant_{i,t}^{CHN}$ represents the migration flow from China to host country i . Thus, the growth rate of the migrant stock is decomposed into a push factor ($\widetilde{Push}_{i,t}$) that encourages labor movement out of China, and a pull factor ($\widetilde{Pull}_{i,t}$) that attracts migration to country i in comparison to other destinations. The former is captured by the growth rate of the global Chinese migrant stock, and the latter by the growth rate of Chinese migrants in i relative to the world. However, to alleviate endogeneity concerns of China-destination-specific factors that drive movements in labor and capital, we employ the “leave-out” approach. Hence, we modify Eq. (5) for the construction of our IV and instead define the push and pull factors as:

$$Push_{i,t} \equiv \frac{\Delta Migrant_{-i,t}^{CHN}}{Migrant_{-i,t-1}^{CHN}}, \text{ and } Pull_{i,t} \equiv \left(\frac{\Delta Migrant_{i,t}^{-CHN}}{Migrant_{i,t-1}^{-CHN}} \right) \div \left(\frac{\Delta Migrant_{WLD,t}^{-CHN}}{Migrant_{WLD,t-1}^{-CHN}} \right).$$

The negative sign of $-i$ denotes all countries other than i . Hence, we associate the push factor to conditions in China but not i , and likewise, the pull factor with i but not China. Furthermore, we follow Burchardi et al. (2019) and rule out plausible sources of endogeneity by interacting $Push_{i,t}$ with $Pull_{i,t}$. Thus, the predicted migrant stock is:

$$\widehat{Migrant}_{i,t}^{CHN} = \begin{cases} Migrant_{i,t-1}^{CHN} \times \left(1 + Push_{i,t} \times Pull_{i,t} \right) & \text{if } Push_{i,t} > 0, \\ Migrant_{i,t-1}^{CHN} \times \left(1 + \frac{Push_{i,t}}{Pull_{i,t}} \right) & \text{otherwise.} \end{cases} \quad (6)$$

¹⁰We found that directly applying the method of Burchardi et al. (2019) here yields extreme predictions of migrant stocks. This is in part due to the decreasing stocks observed in some countries, so that their migrant stocks, which were already small to begin with, turned negative in the predictions.

With negative migration growth rates $Push_{i,t}$ from China, a destination country with a stronger pull factor should lose fewer migrants compared to other destinations, so we divide $Push_{i,t}$ by $Pull_{i,t}$ in this case. We recursively apply Eq. (6) to predict migrant stocks. As with all methods that rely on the shift-share strategy, initial values must be chosen. We set initial values at a 40-year lag by supplementing our data with the World Bank Global Bilateral Migration Database, available from 1960 to 2000 at 10 year intervals.¹¹ For comparison, Javorcik et al. (2011) employ the migrant stock lagged by 30 years as one of the primary instrumental variables in their analysis. With an even longer gap in time between the initial values and our estimation sample, we can also likely rule out reverse causality in which past FDI determines the historical migrant stock.¹² Therefore, the exclusion restriction is plausibly valid given the combination of a lengthy interval between the initial and current values and the variation introduced subsequently through the push and pull factors as described. We use the logarithm of $\widehat{Migrant}^{CHN}$, denoted by $\widehat{Migrant\ network}$, as the instrument for the regressor *Migrant network*.

Using this instrumental variable, we apply the control function approach to estimate both the linear OLS model as well as the nonlinear Poisson and negative binomial models (Wooldridge, 2015). The control function approach is a two-step estimation procedure. In the first step, we regress the endogenous variable *Migrant network* on the instrument $\widehat{Migrant\ network}$ (i.e., the predicted migrant network) and the exogenous variables to obtain the reduced form residuals. In the second step, we regress the number of M&A deals on the exogenous variables and predicted residuals from the first step. It is well known that for linear models, the control function method yields coefficients for the second-stage regressors that are numerically identical to the two-stage least squares (2SLS) estimates (Hausman, 1978; Wooldridge, 2015). However, the control function approach can be adapted to the case where the second step uses a nonlinear model, including the Poisson and negative binomial models.

Table 4 Panel A presents both the first and second-stage estimates. The dependent variable in the first-stage regression is *Migrant network*, the (log) Chinese migrant stock. Because the first stage employs a linear model, the results in columns 1, 3, and 4, which have the same sample size, are identical. They show that our IV, the predicted Chinese migrant network, is significantly correlated with the endogenous variable, the actual Chinese migrant network. Importantly, the second-stage IV results corroborate our main findings for the impact of migrant networks on cross-border M&A. Test diagnostics are reported for the linear models in columns 1 and 2. First, from the Lagrange Multiplier (LM) test for underidentification, the p -values associated with the Kleibergen-Paap rk statistic are less than 0.01, which suggests that the IV is relevant. Next, the weak identification test yields large Cragg-Donald Wald F statistics and Kleibergen-Paap rk Wald F statistics, which reject the null hypothesis that the proposed

¹¹While the UN and World Bank datasets are not identical, the correlation between them is extremely high for the overlapping years of 1990 and 2000 at 0.98.

¹²Although Chinese outward investment only rises significantly at the turn of the 21st century, the overseas Chinese community had been growing long before that, in particular, after restrictions on emigration were lifted in 1860 (Skeldon, 1996). Notable examples of mass migration from China can be traced to labor shortages, including the gold rush in US and Australia, and the development of British colonies in Southeast Asia. Due to such labor market shocks, communities were established, forming the Chinese diasporas observed today.

Table 4: Instrumental Variables Estimates

Panel A: 2SLS and Control Function				
Approach Estimates				
Estimation method	OLS	OLS	Poisson	Negative binomial
Stage I dependent variable	Migrant network			
	(1)	(2)	(3)	(4)
Predicted migrant network	0.193*** (0.048)	0.180*** (0.039)	0.193*** (0.048)	0.193*** (0.048)
Controls and time dummies	Y	Y	Y	Y
Stage II dependent variable	(log) MA+1	(log) MA	MA	MA
Migrant network	0.219*** (0.062)	0.347** (0.149)	0.622*** (0.144)	0.603*** (0.177)
Controls and time dummies	Y	Y	Y	Y
N	943	245	943	943
Shea partial R^2	0.19	0.27		
Underidentification test (p -value)	<0.01	<0.01		
Cragg-Donald Wald F statistic	217.9	80.33		
Kleibergen-Paap rk Wald F statistic	16.48	21.71		
Panel B: CRE2SLS and Control Function				
Approach Estimates				
Estimation method	OLS	OLS	Poisson	Negative binomial
Stage I dependent variable	Migrant network			
	(1)	(2)	(3)	(4)
Predicted migrant network	0.471** (0.235)	0.248 (0.222)	0.471** (0.235)	0.471** (0.235)
Predicted migrant network	-0.245 (0.251)	-0.073 (0.235)	-0.245 (0.251)	-0.245 (0.251)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
Stage II dependent variable	(log) MA+1	(log) MA	MA	MA
Migrant network	0.085 (0.111)	0.803 (0.945)	1.246** (0.538)	1.135*** (0.405)
Predicted migrant network	0.026 (0.026)	-0.098 (0.167)	-0.195 (0.129)	-0.145 (0.096)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
N	943	227	943	943

Notes: The set of control variables is described in the text. Standard errors in parentheses are clustered by host country in columns 1 to 3. ***, **, * denote significance level at 1%, 5% and 10% respectively. p -values associated with the Kleibergen-Paap rk Lagrange Multiplier statistic for the underidentification test are reported.

IV is only weakly related with the endogenous variable. In Appendix Table A.4, we perform an overidentification test using the components of our constructed instrument as separate IVs, namely the (log) 1960 Chinese migrant stock and the interaction of the push and pull factors. The p -values associated with the Hansen J statistics are large and above the threshold of 0.1. Thus, we are unable to reject the null hypothesis that these two instruments are valid and uncorrelated with the error term. This lends support for the validity of the predicted migrant

network as an instrument, since it is constructed from these variables.¹³

In Panel B, we further control for the unobserved heterogeneity of host countries by extending the CRE approach to allow for IV estimation. Hence, this mitigates the concern that destination-specific factors are driving current FDI and the historical migrant stock in 1960. Papke and Wooldridge (2008) and Joshi and Wooldridge (2019) show that the CRE approach works even with endogenous explanatory variables by relying on the the control function approach. Now, in the first step, we regress *Migrant network* on the instrument $\widehat{Migrant\ network}$, the exogenous variables $\mathbf{X}_{1,i,t-1}$ and $\mathbf{X}_{2,i}$ excluding *Migrant network*, and the time averages of all regressors. The residuals from this first stage are obtained. In the second step, we include as regressors in Eq. (4) the first-stage residuals, as well as the time average of the instrument, and estimate the equation using a random effects model. For linear models, the CRE2SLS estimates are also identical to those from a fixed effects 2SLS (i.e., FE2SLS) model (Joshi and Wooldridge, 2019).

In the linear second-stage models of Panel B columns 1 and 2, the impact of the migrant network is positive but imprecisely estimated. In columns 3 and 4, we apply the same control function approach for the nonlinear second-stage model to take into the count nature of the dependent variable. For both the Poisson and negative binomial models, we find that the overseas Chinese migrant network is indeed a strong determinant of the number of cross-border M&A deals from China, even after accounting for the unobserved heterogeneity of host countries. In particular, with the negative binomial model, the migrant network effect is statistically significant at the 1% level.

4.3.2 Greenfield ODI

Besides the IV strategy, we also explore the relation between migrant networks and greenfield ODI to further address the issue of reverse causality. If the positive relationship between migrant networks and cross-border M&A is driven by reverse causality, such that Chinese ODI promotes Chinese migration, then one might expect the same positive relationship with greenfield ODI. As the alternative mode of FDI, greenfield investment involves setting up new production facilities in the host country, and multinationals may send staff from the home country to support these new operations. Therefore, if the direction of causality is reversed, with FDI instead driving migration, we would expect to find it with greenfield investment as its correlation with cross-country labor movements is likely stronger. Such a finding would challenge our previous results. While the absence of a relationship between migrant networks and greenfield ODI cannot prove that there is no reverse causality with respect to cross-border M&A, it would mitigate this concern and imply that Chinese ODI is unlikely to be driving aggregate labor movements abroad.

To investigate this possibility, we obtain transaction-level data for Chinese greenfield ODI from Financial Times Ltd. fDi Markets, and as with the M&A data, aggregate it by host

¹³We are also unable to reject the null hypothesis of the overidentification test when the predicted migrant network and, for example, the destination pull factor are used as instruments.

Table 5: Overseas Chinese Migrant Networks and Greenfield ODI

Estimation method	Negative binomial			Predicted migrant network
	(1)	(2)	(3)	
Instrument				
Dependent variable		GF		
	(1)	(2)	(3)	(4)
Migrant network	0.053 (0.046)		0.070 (0.044)	-0.246 (0.232)
Chinese share of migrants		0.009 (0.112)		
Predicted migrant network				0.054 (0.059)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
Industry dummies	N	N	Y	N
<i>N</i>	966	966	9,660	943

Notes: The set of control variables is described in the text. Column 4 presents the second-stage regression results with the predicted migrant network as the instrument using the control function approach. ***, **, * denote significance level at 1%, 5% and 10% respectively.

country for the dependent variable.¹⁴ Moreover, we restrict the analysis to the same set of host countries for comparability to the previous findings. For the remainder of this paper, results from the negative binomial regression are presented. However, similar conclusions are drawn with the Poisson model as shown in Appendix Table A.5.

Table 5 shows that, across all columns, the relationship between migrant networks and greenfield ODI is insignificant and economically indistinguishable from zero. Column 1 shows the baseline result, while column 2 demonstrates robustness using the alternative measure of the Chinese share of migrants. Column 3 presents the host country-industry level regression, and column 4 employs the same IV strategy as above. Thus, we find no evidence from Table 5 that Chinese migrant networks impact Chinese greenfield ODI. This suggests that reverse causality is not a strong concern. If past investment drives both current investment and migration, then a positive relationship should be observed between the presence of migrant networks and all modes of FDI, not M&A alone. However, no such relationship is found with greenfield investment from China.

In theory, migrant networks may also play a role in facilitating greenfield investment. However, we do not find empirical evidence of this in the case of Chinese greenfield ODI. Even though both modes of FDI are conducted to gain access to a foreign market, the entry barriers that greenfield and M&A FDI face can be quite different. Hence, this may lead to the heterogeneous roles of migrant networks on FDI. While both greenfield investment and cross-border M&A allow firms to expand across national boundaries, the latter differs in that it also crosses organization boundaries. Because M&A involve the transfer of ownership, there are many problems associated with negotiation, bargaining, and contract enforcement that are not present in

¹⁴UNCTAD (2017) and Davies et al. (2018) use the same dataset for analyzing global greenfield investment.

greenfield investment. In addition to the costs of searching for an appropriate target, sealing a cross-border M&A deal requires skills in communication and execution, and a sophisticated understanding and information about the capabilities, preferences, and potential synergies of both the target's and the acquirer's businesses. After the deal is completed, daily business operations also benefit greatly from knowledge about the local markets. For instance, connections to local suppliers can be important and understanding local preferences is key for marketing. Migrants are better at extracting relevant information through their networks or connections with local investors and processing them efficiently, which help multinational investors overcome the information barriers. In these aspects, cross-border M&A requires more intensive information than greenfield investment and may therefore rely more on migrant networks. Furthermore, while greenfield investment creates job opportunities and technology spillovers at new production facilities, cross-border M&A tends to be associated with little job creation and absorbs technology from acquired companies. Hence, host countries generally welcome greenfield investment more than M&A (UNCTAD, 2014). Governments may also restrict foreign acquisitions due to concerns of national security, job loss, and the protection of firms in strategic industries (UNCTAD, 2017; Bertrand et al., 2012), which poses higher barriers for cross-border M&A.

Davies et al. (2018) study the determinants of greenfield and M&A FDI jointly, and their empirical evidence is supportive of our finding for the heterogeneous migrant network effect across entry modes. Indeed, they make similar arguments that as a contract intensive transaction, cross-border M&A is affected more by barriers and frictions in the destination. In particular, Davies et al. (2018) hypothesize that M&A is more deterred by international barriers, whether geographic, cultural, or policy-driven. Moreover, compared to greenfield investment, M&A is more dependent on the institutional quality and financial development of the destination. Using global M&A and greenfield FDI data, Davies et al. (2018) find strong empirical evidence in support of their hypotheses. Importantly, their results are consistent with our findings for the positive impact of migrant networks on cross-border M&A, but not greenfield investment. As we demonstrate next in Section 5, many of the barriers that Davies et al. (2018) examine are exactly the ones which are alleviated by the overseas migrant networks.

5 Heterogeneity analysis

In this section, we investigate the migrant network effect on cross-border M&A in a dynamic context by examining the change in the role of migrants as MNEs learn from their investment experience. We also study the heterogeneous roles of migrant networks in dealing with implicit and explicit information barriers by exploiting the various dimensions of information frictions at the firm, industry, and country levels. Results from the CRE negative binomial model are reported, but the Poisson model delivers similar findings, as shown in Appendix Table A.6.

Table 6: Heterogeneity by Overseas Investment Experience

Estimation method Dependent variable	Negative binomial			
	MA			
	(1)	(2)	(3)	(4)
Migrant network	0.228*** (0.072)	0.173** (0.076)	0.039 (0.078)	0.221*** (0.073)
First	0.868*** (0.086)	-0.140 (0.483)		
Migrant network \times First		0.085** (0.040)		
(log) Cumulative MA			-0.702*** (0.186)	-1.053** (0.481)
(log) Cumulative MA			1.391*** (0.239)	3.400*** (0.560)
Migrant network \times (log) Cumulative MA				0.032 (0.038)
Migrant network \times (log) Cumulative MA				-0.166*** (0.042)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
N	1,932	1,932	966	966

Notes: The set of control variables is described in the text. The constant 1 is added to Cumulative MA before taking the logarithm. ***, **, * denote significance level at 1%, 5% and 10% respectively.

5.1 Overseas investment experience

One channel through which migrant networks may facilitate ODI is the alleviation of information frictions. If there are local knowledge spillovers from social networks to multinationals, the role of migrant networks in promoting cross-border investment should decline as MNEs accumulate knowledge and experience overseas. Thus, as in Section 2, we now classify M&A transactions into two groups based on the investor’s overseas investment experience. One group is comprised of acquirers’ first cross-border M&A deal, while the second group includes all subsequent deals abroad. We are able to decompose our sample in this way because of the availability of cross-border M&A data since 1981. The annual investment of Chinese MNEs to each destination is then aggregated for first and subsequent entry separately. Since the former group is less experienced, the information barriers encountered would tend to be stronger, and the influence of migrants is expected to be more pronounced if they do play a role in alleviating information frictions.

In Table 6 columns 1 and 2, we generate a dummy variable *First* that equals to 1 for the group of first entrants. Thus, analogous to the industry-level analysis in Table 3 Panel C, we now define “level 1” variables at the host country-group level. Note that the time average of this dummy variable, \overline{First} , is exactly equal to the variable itself, so it is dropped from the regression. Because the time-varying covariates in $\mathbf{X}_{1,i,t-1}$ do not vary along the group dimension, their time averages $\overline{\mathbf{X}}_{1,i,t-1}$ are identical to before. In column 1, we control for the varying composition of first and subsequent deals. The coefficient of the dummy variable *First* is positive, consistent with our stylized fact that the majority of acquisitions is originating from

new investors.

In column 2, we interact *Migrant network* with *First*. Because we have two groups (i.e., *First* is either 0 or 1), the time average of a time-invariant variable interacted with *First* is exactly half of the variable itself. This means that the interaction term is perfectly collinear with the time-invariant variable, so it is dropped from the regression. Given that we treat the migrant network as time-invariant, we also exclude the time average of $Migrant\ network \times First$. In column 2, we find a more pronounced migrant network effect for the first-time entrants. This is consistent with our prediction that knowledge transfers from migrant networks to MNEs promote greater cross-border investment due to the lack of experience by the investors. Moreover, these results are unlikely to be driven by the size of M&A deal, as the average values of first and subsequent deals are very close at 305 and 314 million USD, respectively.

As we have demonstrated, knowledge spillovers may occur as MNEs interact with social networks. At the same time, knowledge about the foreign destination may also be disseminated domestically from investing companies to future investors within China. For instance, this may be achieved through informal channels such as business networks or media coverage. Hence, the cumulative overseas investment experience by Chinese investors is expected to lower the information frictions encountered by MNEs, and the migrant network effect may diminish as the investors become more experienced.

We test this hypothesis by controlling for cumulative Chinese M&A in a host country as a proxy for the overall experience of Chinese investors in the destination. Because this covariate is time varying, we must include its time average. The regression results in column 3 present both the within and between estimates. In particular, we find that the coefficient of the time average $\overline{(\log) Cumulative MA}$ is positive and statistically significant. This between estimate exploits cross-sectional variation, and indicates that cross-border M&A deals are directed towards host countries for which China as a whole has more experience. This supports the idea of knowledge spillovers across MNEs. The results show that the within estimate is negative, suggesting that the spillovers towards a given host country have a diminishing impact over time. In column 4, the coefficient on the interaction term $\overline{Migrant\ network \times (\log) Cumulative MA}$ is negative. This further implies that migrant networks serve as a substitute source of knowledge spillovers from experienced Chinese investors in M&A.

5.2 Heterogeneity across firms and industries

As with the exchange of goods and services, information asymmetry is a problem for the transfer of ownership between two parties, especially when the transaction is made across international borders. Moreover, the intensity of information required is perhaps greater than other forms of transactions, as the acquirer and target must have a mutual understanding of the culture and management style of each firm to generate greater value and synergies. The deal-level M&A data provides some information on firms' characteristics in the years for which M&A activity is observed. Therefore, we now exploit variation in firm characteristics to further understand the role of migrant networks in mitigating firm-level information barriers.

Table 7: Firm and Industry Heterogeneity

Estimation method Dependent variable	Negative binomial					
	MA					
	Unlisted acquirer or target	Listed acquirer and target	Private- owned enterprise	State- owned enterprise	Across industries	Within industries
(1)	(2)	(3)	(4)	(5)	(6)	
Migrant network	0.164** (0.077)	0.524** (0.243)	0.377*** (0.094)	-0.069 (0.123)	0.164* (0.091)	0.173 (0.115)
First	-0.162 (0.495)	0.824 (1.945)	1.815*** (0.673)	-2.147*** (0.795)	-0.254 (0.605)	0.027 (0.784)
Migrant network \times First	0.089** (0.041)	-0.042 (0.161)	-0.060 (0.054)	0.227*** (0.071)	0.107** (0.051)	0.050 (0.064)
Controls and time dummies	Y	Y	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y	Y	Y
<i>N</i>	1,932	1,932	1,932	1,932	1,932	1,932

Notes: The set of control variables is described in the text. ***, **, * denote significance level at 1%, 5% and 10% respectively.

First, we distinguish firms by their stock exchange listing status. Companies listed on a stock exchange have more information about them revealed and accessible to the general public. In particular, they must disclose certain financial details and file earnings reports, enhancing the transparency of their operations (Erel et al., 2012). Thus, we expect migrants to be more important in facilitating transactions between unlisted acquirer and target firms, and especially when they have less experience abroad. To test this hypothesis, we follow Erel et al. (2012) and use the information on a firm’s “public” status to identify whether it is listed on a stock exchange. We then count the number of M&A deals from China for two subsamples: (i) unlisted acquirers or targets, and (ii) listed acquirers and targets. For each subsample, we also distinguish between the acquirer’s first and all subsequent acquisitions abroad.

The results are reported in Table 7. In column 1, for unlisted acquirers or targets, we find a positive and statistically significant coefficient on the interaction term *Migrant network* \times *First*. We do not find this effect in column 2 for listed companies. This is consistent with the idea that unlisted companies are less transparent. Hence, there is a greater role for migrant networks in alleviating firm-specific information barriers for new entrants into the host country.

Next, we explore the potential heterogeneity across firm ownership structure. For each acquiring company, we collect information on its ownership structure from Orbis, a database of 200 million companies worldwide managed by Bureau van Dijk. A firm is classified as a state-owned enterprises (SOE) if at least 25.01% of its ownership belongs ultimately to the Chinese government, and a privately-owned enterprises (POE) otherwise.¹⁵ *A priori*, whether the impact of migrants is greater for SOEs or POEs is not obvious. On the one hand, SOEs have better access to resources, such as the assistance of embassies or other government entities, than POEs. On the other hand, POEs are found to be more active in utilizing overseas Chinese networks compared to SOEs. According to a survey by the China Council for the Promotion of

¹⁵We confirm our regression results are also robust when we change the equity threshold of the global ultimate owner to 50.01%.

International Trade, 51.1% POEs turn to the overseas Chinese Chamber of Commerce in dealing with investment risk, while only 38.6% for SOEs ask for this assistance (CCPIT, 2015, Figure 6.11). The survey also reports that 93.5% POEs highly value opinions from their overseas employees as opposed to 80.7% for SOEs (CCPIT, 2015, Figure 6.8). Therefore, in order to see which effect dominates, we aggregate the number of M&A deals by ownership type to repeat the analysis above. The results are presented in Table 7 columns 3 and 4. Empirically, the overseas Chinese migrants have a stronger impact on the SOEs that are less experienced. This suggests that SOEs that enter a host country for the first time benefit more from the assistance of migrant networks, possibly due to their connections to various government entities.

If a deal is made outside of the acquirer’s own industry, the firm may also find the environment less familiar, and encounter greater information asymmetry compared to those that purchase a target in the same industry. Thus, the migrant network effect is expected to be greater when acquirers and targets are operating in different sectors. Cross-border M&A transactions are classified as “across industries” if the 2-digit industry code of the acquirer and target companies differ. The results in Table 7 columns 5 and 6 show that the role of migrant networks is indeed more pronounced for M&A across industries for their first entry. We do not find this effect for first-time entrants purchasing a target within the same industry.

5.3 Heterogeneity across host countries

Information asymmetries associated with individual firms and industries may be compounded by the barriers encountered in a host country when dealing with business regulations and restrictions. Again, the MNE’s decision to enter a foreign country will be influenced by the degree to which the economic environment and frictions in factor markets impede operations and raise costs. Given that migrants possess an information advantage, we expect them to have a greater impact in helping inexperienced firms navigate the regulatory environment in host countries with stricter barriers. From the Economic Freedom of the World database (Gwartney et al., 2015), we obtain six measures of country-level market frictions and regulatory barriers. The definitions of these variables are listed in Appendix A.1. We follow the strategy from before, and separate the cross-border M&A deals from China into two subsamples of host countries. To contrast those with strict versus weak regulations, each of the six variables is converted to an indicator variable that takes the value of 1 if the measure of the country’s barrier is greater than or equal to the median value across the sample, and zero otherwise. For each subsample, we examine the interaction between the size of the migrant network and the dummy variable for first entry.

The findings for each of the six measures of regulatory obstacles are presented in Table 8. Specifically, these measures are: (i) overall regulations, (ii) overall business regulations, (iii) the prevalence of extra payments or bribes and favoritism, (iv) hiring and firing regulations in the labor market, (v) overall credit market regulations, and (vi) foreign ownership and investment restrictions. While measure (v) captures general capital market frictions, measure (vi) pertains specifically to cross-border investment. For each measure, the odd (even) column shows the

Table 8: Host-Country Heterogeneity in Regulatory Barriers

Estimation method Dependent variable	Panel A					
	Negative binomial					
	MA					
	Regulation index		Business regulations		Bribes, favoritism	
High	Low	High	Low	High	Low	
(1)	(2)	(3)	(4)	(5)	(6)	
Migrant network	0.063 (0.104)	0.568*** (0.090)	-0.117 (0.109)	0.415*** (0.110)	-0.055 (0.140)	0.404*** (0.096)
First	-0.945 (0.877)	0.865 (0.647)	-1.518* (0.811)	1.164* (0.659)	-1.610* (0.954)	1.039 (0.684)
Migrant network \times First	0.152* (0.087)	0.008 (0.051)	0.212*** (0.079)	-0.017 (0.052)	0.213** (0.097)	-0.004 (0.054)
Controls and time dummies	Y	Y	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y	Y	Y
<i>N</i>	942	924	928	910	886	886

Estimation method Dependent variable	Panel B					
	Negative binomial					
	MA					
	Hiring and firing regulations		Credit market regulations		Foreign ownership restrictions	
High	Low	High	Low	High	Low	
(1)	(2)	(3)	(4)	(5)	(6)	
Migrant network	0.339** (0.138)	-0.0004 (0.093)	0.065 (0.110)	0.313*** (0.111)	-0.008 (0.100)	0.373*** (0.116)
First	-1.111 (0.862)	0.055 (0.625)	-0.763 (0.692)	0.675 (0.672)	-0.291 (0.618)	0.203 (0.766)
Migrant network \times First	0.181** (0.076)	0.063 (0.050)	0.128** (0.060)	0.025 (0.054)	0.092* (0.051)	0.062 (0.063)
Controls and time dummies	Y	Y	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y	Y	Y
<i>N</i>	894	878	936	930	894	878

Notes: The set of control variables is described in the text. ***, **, * denote significance level at 1%, 5% and 10% respectively.

results for the subsample with high (low) regulatory barriers. In both Panels A and B, we find that the coefficient on the interaction term *Migrant network* \times *First* is positive and statistically significant in the odd columns, but in not the even columns. Hence, the results imply that for inexperienced investors, the Chinese migrant networks are able to facilitate entry into countries with burdensome government bureaucracy and highly regulated factor markets. In contrast, the migrant network effect is not observed for first-time entrants when regulations are loose and labor and capital market frictions are low. In summary, Table 8 is consistent with overseas migrant networks providing the necessary information and local knowledge to help navigate regulatory barriers and overcome various market frictions, especially for first-time entrants.

6 Conclusion

Using data on Chinese migrant stocks and FDI across many countries, this paper provides empirical evidence that migrant networks promote cross-border M&A. We show a stronger

impact of overseas Chinese migrant networks in environments with greater information frictions. The relationship between migrant networks and ODI is more pronounced for less experienced investors, especially between non-listed firms, across industries, and in host countries with higher regulatory barriers.

The growth of international migrant stocks and investment is not unique to China alone. In fact, the largest international migrant stocks all originate from developing countries, namely, India, Mexico, Russia, China, and Bangladesh. With rising incomes in developing countries and greater labor mobility, the global stock of migrants continues to rise. Moreover, this has been accompanied by tremendous growth in cross-border capital flows: the outward stock of FDI (as a percentage of GDP) for the world has tripled since 1990 (from 10.1% to 34.0%).¹⁶ For countries like India and South Africa, FDI outflows are 2.5 and 6 times larger, respectively, since 2005. Thus, understanding the linkages in factors of production is key to the future research agenda.

¹⁶See Annex Table 8 from (UNCTAD, 2017).

References

- Abel, G. J. and J. E. Cohen (2019). Bilateral international migration flow estimates for 200 countries. *Scientific Data* 6, Article 82.
- Aizenman, J., Y. Jinjarak, and H. Zheng (2018). Chinese outwards mercantilism – the art and practice of bundling. *Journal of International Money and Finance* 86, 31 – 49.
- Allison, P. D. (2005). *Fixed effects regression methods for longitudinal data using SAS*. Cary, NC: SAS Institute Inc.
- Azemar, C., J. Darby, R. Desbordes, and I. Wooton (2012). Market familiarity and the location of South and North MNEs. *Economics & Politics* 24 (3), 307–345.
- Bartik, T. J. (1991). Who benefits from state and local economic development policies? In *Books from Upjohn Press, W.E. Upjohn Institute for Employment Research, number wbsle, November*.
- Bekaert, G., C. R. Harvey, and C. Lundblad (2004). Does financial liberalization spur growth? *Journal of Financial Economics* 77, 3–55.
- Bénassy-Quéré, A., M. Coupet, and T. Mayer (2007). Institutional determinants of foreign direct investment. *World Economy* 30(5), 764–782.
- Bertrand, O., Hakkala, K. Nilsson, P.-J. Norbäck, and L. Persson (2012). Should countries block foreign takeovers of R&D champions and promote greenfield entry. *Canadian Journal of Economics* 45(3), 1083–1124.
- Bhattacharya, U. and P. Groznik (2008). Melting pot or salad bowl: Some evidence from U.S. investments abroad. *Journal of Financial Markets* 11(3), 228–258.
- Bisztray, M., M. Koren, and A. Szeidl (2018). Learning to import from your peers. *Journal of International Economics* 115, 242–258.
- Blonigen, B. A. and D. Lee (2016). Heterogeneous frictional costs across industries in cross-border mergers and acquisitions. *NBER Working Papers* 22546.
- Blonigen, B. A. and J. Piger (2014). Determinants of foreign direct investment. *Canadian Journal of Economics* 47(3), 775–812.
- Buch, C. M., J. Kleinert, and F. Toubal (2006). Where enterprises lead, people follow? Links between migration and FDI in Germany. *European Economic Review* 50(8), 2017–2036.
- Burchardi, K. B., T. Chaney, and T. A. Hassan (2019). Migrants, ancestors, and foreign investments. *Review of Economic Studies* 86(4), 1448–1486.
- Burchardi, K. B. and T. A. Hassan (2013). The economic impact of social ties: Evidence from German reunification. *Quarterly Journal of Economics* 128(3), 1219–1271.

- Cameron, A. C. and P. K. Trivedi (2009). *Microeconometrics using Stata*. College Station, Texas: StataCorp LP.
- Cameron, A. C. and P. K. Trivedi (2013). *Regression Analysis of Count Data* (2 ed.). Cambridge, MA: Cambridge University Press.
- Chaney, T. (2014). The network structure of international trade. *American Economic Review* 104(11), 3600–3634.
- Chang, P.-I. (2014). Complementarity in institutional quality in bilateral FDI flows. *SMU Economics & Statistics Working Paper Series, Paper No. 20-2014*.
- Chen, C., W. Tian, and M. Yu (2019). Outward FDI and domestic input distortions: Evidence from Chinese firms. *Economic Journal* 129, 3025–3057.
- Chen, W. and H. Tang (2014). The dragon is flying west: Micro-level evidence of Chinese outward direct investment. *Asian Development Review* 31(2), 109–140.
- Cheung, Y.-W. and X. Qian (2009). Empirics of China’s outward direct investment. *Pacific Economic Review* 14(3), 312–641.
- China Council for the Promotion of International Trade (CCPIT) (2015). Survey on Chinese enterprises’ outbound investment and operation in 2012. *China Council for the Promotion of International Trade*.
- Cohen, L., U. G. Gurun, and C. Malloy (2017). Resident networks and corporate connections: Evidence from World War II internment camps. *Journal of Finance* 72(1), 207–248.
- Combes, P.-P., M. Lafourcadea, and T. Mayer (2005). The trade-creating effects of business and social networks: Evidence from France. *Journal of International Economics* 66(1), 1–29.
- Davies, R. B., R. Desbordes, and A. Ray (2018). Greenfield versus merger & acquisition FDI: Same wine, different bottles? *Canadian Journal of Economics* 51(4), 1151–1190.
- de Sousa, J. (2012). The currency union effect on trade is decreasing over time. *Economics Letters* 117(3), 917–920.
- Desbordes, R. and S.-J. Wei (2017). The effects of financial development on foreign direct investment. *Journal of International Economics* 127, 153–168.
- di Giovanni, J. (2005). What drives capital flows? The case of cross-border M&A activity and financial deepening. *Journal of International Economics* 65, 127–149.
- Dunlevy, J. A. (2006). The influence of corruption and language on the protrade effect of immigrants: Evidence from the American states. *Review of Economics and Statistics* 88(1), 182–186.

- Erel, I., R. C. Liao, and M. S. Weisbach (2012). Determinants of cross-border mergers and acquisitions. *Journal of Finance* 67(3), 1045–1082.
- Fernandes, A. P. and H. Tang (2014). Learning to export from neighbors. *Journal of International Economics* 94(1), 67 – 84.
- Gao, T. (2003). Ethnic Chinese networks and international investment: Evidence from inward FDI in China. *Journal of Asian Economics* 14(4), 611–629.
- Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review* 110(8), 2586–2624.
- Gould, D. M. (1994). Immigrant links to the home country: Empirical implications for U.S. bilateral trade flows. *Review of Economics and Statistics* 76(2), 302–316.
- Greene, W. H. (2012). *Econometric Analysis* (7 ed.). Upper Saddle River, NJ: Prentice Hall.
- Gwartney, J., R. Lawson, and J. Hall (2015). 2015 Economic Freedom Dataset, published in Economic Freedom of the World: 2015 Annual Report. http://www.freetheworld.com/datasets_efw.html.
- Hausman, J. (1978). Specification tests in econometrics. *Econometrica* 46(6), 1251–1271.
- Head, K. and J. Ries (1998). Immigration and trade creation: Econometric evidence from Canada. *Canadian Journal of Economics* 31(1), 47–62.
- Head, K. and J. Ries (2008). FDI as an outcome of the market for corporate control: Theory and evidence. *Journal of International Economics* 74, 2–20.
- Huang, J. and D. J. Kisgen (2013). Gender and corporate finance: Are male executives overconfident relative to female executives? *Journal of Financial Economics* 108(3), 822–839.
- Huang, Y., L. Jin, and Y. Qian (2013). Does ethnicity pay? Evidence from overseas Chinese FDI in China. *Review of Economics and Statistics* 95(3), 868–883.
- Huang, Y. and B. Wang (2013). Investing overseas without moving factories abroad: The case of Chinese outward direct investment. *Asian Development Review* 30(1), 85–107.
- Javorcik, B. S., Ç. Özden, M. Spatareanu, and C. Neagu (2011). Migrant networks and foreign direct investment. *Journal of Development Economics* 94, 231–241.
- Joshi, R. and J. M. Wooldridge (2019). Correlated random effects models with endogenous explanatory variables and unbalanced panels. *Annals of Economics and Statistics* 134, 243–268.
- Kamal, F. and A. Sundaram (2016). Buyer–seller relationships in international trade: Do your neighbors matter? *Journal of International Economics* 102, 128 – 140.

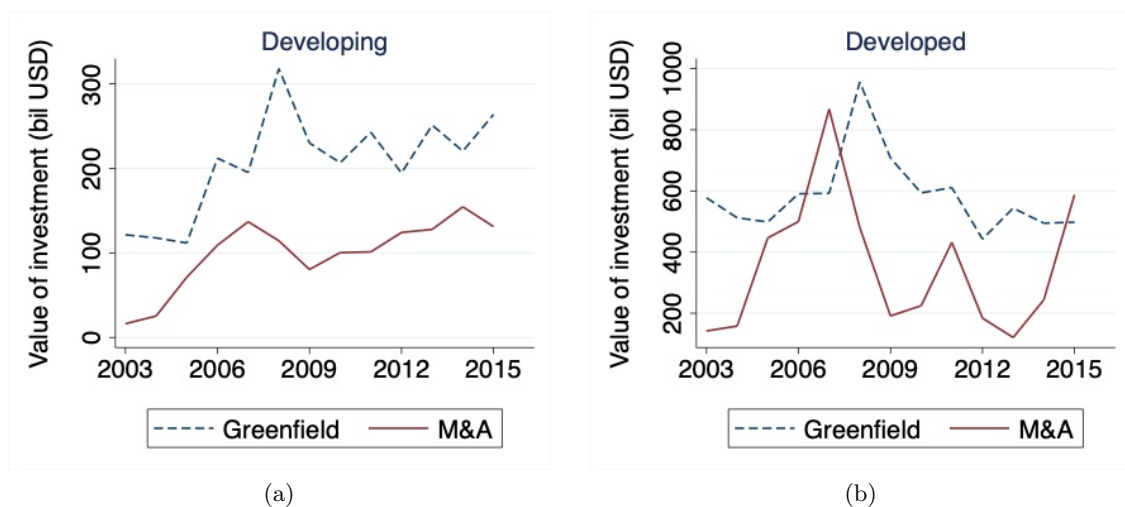
- Kugler, M. and H. Rapoport (2007). International labor and capital flows: Complements or substitutes? *Economic Letters* 94(2), 155–162.
- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica* 46(1), 69–85.
- Ottaviano, G. I., G. Peri, and G. C. Wright (2018). Immigration, trade and productivity in services: Evidence from U.K. firms. *Journal of International Economics* 112, 88 – 108.
- Papke, L. E. and J. M. Wooldridge (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics* 145(1-2), 121–133.
- Parsons, C. and P.-L. Vézina (2018). Migrant networks and trade: The Vietnamese boat people as a natural experiment. *Economic Journal* 128(612), 210–234.
- Rauch, J. E. (2001). Business and social networks in international trade. *Journal of Economic Literature* 39(4), 1177–1203.
- Rauch, J. E. and V. Trindade (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics* 84(1), 116–130.
- Rossi, S. and P. F. Volpin (2004). Cross-country determinants of mergers and acquisitions. *Journal of Financial Economics* 74, 277–304.
- Skeldon, R. (1996). Migration from China. *Journal of International Affairs* 49(2), 434–455.
- Steingress, W. (2018). The causal impact of migration on US trade: Evidence from a natural experiment. *Canadian Journal of Economics* 51(4), 1312–1338.
- Tong, S. Y. (2005). Ethnic networks in FDI and the impact of institutional development. *Review of Development Economics* 9(4), 563–580.
- UNCTAD (2014). *World investment report*. New York.
- UNCTAD (2017). *World investment report*. New York.
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2 ed.). Cambridge, MA: MIT Press.
- Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources* 50(2), 420–445.
- Wooldridge, J. M. (2019). Correlated random effects models with unbalanced panels. *Journal of Econometrics* 211, 137–150.

A Appendix

A.1 Definitions of regulatory barriers (Gwartney et al., 2015)

This database covers five areas of economic freedom: (1) Size of government, (2) Legal system and property rights, (3) Sound money, (4) Freedom to trade internationally, (5) Regulation. Our analysis utilizes the following categories related to regulatory barriers and restrictions:

- Regulation index (Area 5): An average index is constructed from the indices of: (A) Credit market regulations, (B) Labor market regulations, and (C) Business regulations.
- Business regulations (component of Area 5 (Regulation)): An average index is constructed from the indices of 6 subcomponents: (i) Administrative requirements, (ii) Bureaucracy costs, (iii) Starting a business, (iv) Extra payments/bribes/favoritism, (v) Licensing restrictions. (vi) Cost of tax compliance.
- Bribes, favoritism (subcomponent (C)(iv) of Area 5 (Regulation)): Based on the Global Competitiveness Report questions: (1) “In your industry, how commonly would you estimate that firms make undocumented extra payments or bribes connected with the following: Import and export permits; Connection to public utilities (e.g., telephone or electricity); Annual tax payments; Awarding of public contracts (investment projects); Getting favorable judicial decisions. Common (= 1), Never occur (= 7)”. (2) “Do illegal payments aimed at influencing government policies, laws or regulations have an impact on companies in your country? 1 = Yes, significant negative impact, 7 = No, no impact at all”. (3) “To what extent do government officials in your country show favoritism to well-connected firms and individuals when deciding upon policies and contracts? 1 = Always show favoritism, 7 = Never show favoritism”.
- Hiring and firing regulations (subcomponent (B)(ii) of Area 5 (Regulation)): Based on the Global Competitiveness Report question: “The hiring and firing of workers is impeded by regulations (= 1) or flexibly determined by employers (= 7)”.
- Credit market regulations (component of Area 5 (Regulation)): An average index is constructed from the indices of 3 subcomponents: (i) Ownership of banks, (ii) Private sector credit, (iii) Interest rate controls/negative real interest rates.
- Foreign ownership restrictions (subcomponent (D)(i) of Area 4 (Freedom to trade internationally)): Based on the following questions from the Global Competitiveness Report: (1) “How prevalent is foreign ownership of companies in your country? 1 = Very rare , 7 = Highly prevalent”; (2) “How restrictive are regulations in your country relating to international capital flows? 1 = Highly restrictive, 7 = Not restrictive at all.”



Appendix Figure A.1: Outward greenfield FDI and M&A from (a) developing countries and transition economies; (b) developed economies. Source: UNCTAD (2017).

Appendix Table A.1: Summary Statistics

Variable	Mean	Median	Std Dev	Minimum	Maximum
MA	1.22	0	4.08	0	37
Migrant network	8.28	7.82	2.41	4.33	14.6
(log) GDP	25.3	25.3	1.98	20.4	30.5
(log) GDP per capita	8.88	8.99	1.47	5.49	11.6
(log) Population	9.48	9.40	1.69	5.66	14.1
(log) Distance	9.00	9.03	0.58	7.06	9.86
Common language	0.04	0	0.19	0	1
Free trade agreement	0.09	0	0.29	0	1
Trade openness	0.92	0.75	0.65	0.21	4.43
Natural resources	0.26	0.15	0.26	0.003	0.99
GDP growth rate	3.77	3.89	3.80	-14.8	26.3
Financial development	0.70	0.47	0.57	0.0001	3.14
Institutional quality	9.35	8.50	3.20	3	16
Chinese migrant share	0.58	0.03	3.41	0.0002	32.3
GF	2.69	0	7.11	0	80

Appendix Table A.2: Full Results from CRE Model of Table 2 Panel A

Estimation method Dependent variable	OLS	OLS	Poisson	Negative
	(log) MA+1 (1)	(log) MA (2)	MA (3)	binomial MA (4)
Migrant network	0.072*** (0.025)	0.137* (0.073)	0.176 (0.275)	0.233*** (0.070)
(log) GDP	-1.673 (1.203)	-13.942** (6.865)	-2.424 (5.800)	-3.737 (6.131)
(log) GDP per capita	1.390 (1.217)	13.269* (6.774)	2.387 (5.699)	3.708 (6.094)
(log) Population	0.095 (1.316)	14.511 (9.219)	-0.538 (7.065)	0.767 (6.662)
(log) Distance	0.072 (0.070)	0.160 (0.127)	0.138 (0.151)	0.248 (0.155)
Common language	0.619 (0.395)	0.241 (0.333)	0.117 (0.333)	0.217 (0.608)
Free trade agreement	-0.083 (0.099)	-0.240 (0.307)	-0.617 (0.395)	-0.456 (0.346)
Trade openness	-0.111 (0.125)	-0.442 (0.615)	0.616 (0.520)	0.625 (0.521)
Natural resources	-0.122 (0.286)	1.096 (1.896)	-1.172 (1.746)	-1.549 (1.413)
GDP growth rate	0.005 (0.004)	0.003 (0.031)	0.024 (0.029)	0.013 (0.027)
Financial development	0.037 (0.102)	-0.647 (0.462)	0.015 (0.473)	-0.043 (0.392)
Institutional quality	0.034 (0.036)	-0.143 (0.118)	-0.129 (0.165)	-0.136 (0.108)
<u>(log) GDP</u>	3.384*** (1.265)	16.051** (7.307)	3.841 (6.645)	4.983 (7.090)
<u>(log) GDP per capita</u>	-3.035** (1.273)	-15.193** (7.209)	-2.962 (6.600)	-4.075 (7.073)
<u>(log) Population</u>	-1.704 (1.353)	-16.537* (9.696)	-0.107 (7.998)	-1.390 (7.562)
<u>Free trade agreement</u>	-0.136 (0.192)	0.061 (0.530)	1.209 (0.852)	0.953 (0.634)
<u>Trade openness</u>	0.082 (0.144)	0.429 (0.608)	-0.334 (0.524)	-0.462 (0.597)
<u>Natural resources</u>	0.350 (0.242)	-0.521 (1.754)	3.108* (1.600)	3.126** (1.523)
<u>GDP growth rate</u>	0.033 (0.021)	0.075 (0.069)	0.197 (0.201)	0.265*** (0.095)
<u>Financial development</u>	0.265 (0.203)	0.800 (0.518)	0.608 (0.863)	0.598 (0.499)
<u>Institutional quality</u>	-0.018 (0.047)	0.165 (0.145)	0.191 (0.198)	0.174 (0.124)
Time dummies	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	966	248	966	966

Notes: Standard errors in parentheses are clustered by host country in columns 1 to 3. ***, **, * denote significance level at 1%, 5% and 10% respectively.

Appendix Table A.3: CRE Model Estimates Treating *Migrant Network* as Time-Varying

Estimation method Dependent variable	OLS	OLS	Poisson	Negative
	(log) MA+1 (1)	(log) MA (2)	MA (3)	binomial MA (4)
Migrant network	-0.013 (0.032)	-0.165 (0.121)	-0.086 (0.150)	-0.031 (0.148)
<u>Migrant network</u>	0.136*** (0.042)	0.363*** (0.129)	0.357** (0.151)	0.316* (0.164)
(log) GDP	-1.602 (1.152)	-11.018 (7.506)	-1.860 (5.588)	-2.993 (6.044)
(log) GDP per capita	1.291 (1.165)	10.383 (7.382)	1.877 (5.470)	2.996 (6.006)
(log) Population	-0.065 (1.259)	10.573 (9.322)	-1.218 (6.635)	0.082 (6.546)
(log) Distance	0.093 (0.069)	0.188 (0.131)	0.170 (0.142)	0.258* (0.155)
Common language	0.513 (0.346)	0.238 (0.322)	-0.064 (0.495)	0.044 (0.621)
Free trade agreement	-0.068 (0.084)	-0.172 (0.215)	-0.638** (0.321)	-0.520 (0.344)
Trade openness	-0.110 (0.129)	-0.067 (0.677)	0.782 (0.518)	0.769 (0.522)
Natural resources	-0.117 (0.288)	1.183 (2.053)	-1.068 (1.676)	-1.408 (1.404)
GDP growth rate	0.006 (0.004)	-0.001 (0.032)	0.023 (0.031)	0.013 (0.027)
Financial development	0.078 (0.103)	-0.579 (0.500)	0.187 (0.469)	0.131 (0.403)
Institutional quality	0.027 (0.035)	-0.202*** (0.076)	-0.138 (0.159)	-0.136 (0.108)
<u>(log) GDP</u>	3.181*** (1.210)	12.853 (7.923)	3.712 (6.723)	4.627 (7.022)
<u>(log) GDP per capita</u>	-2.828** (1.220)	-12.141 (7.804)	-2.985 (6.620)	-3.798 (7.003)
<u>(log) Population</u>	-1.455 (1.296)	-12.375 (9.798)	0.031 (7.910)	-1.153 (7.462)
<u>Free trade agreement</u>	-0.151 (0.184)	-0.115 (0.385)	1.247 (0.766)	1.084* (0.625)
<u>Trade openness</u>	0.080 (0.142)	0.066 (0.691)	-0.509 (0.568)	-0.606 (0.604)
<u>Natural resources</u>	0.326 (0.256)	-0.639 (1.901)	2.857* (1.602)	2.888* (1.518)
<u>GDP growth rate</u>	0.031 (0.020)	0.065 (0.068)	0.184 (0.174)	0.258*** (0.095)
<u>Financial development</u>	0.159 (0.194)	0.676 (0.546)	0.370 (0.747)	0.407 (0.510)
<u>Institutional quality</u>	-0.011 (0.045)	0.230** (0.103)	0.189 (0.190)	0.162 (0.124)
Time dummies	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
<i>N</i>	966	248	966	966

Notes: Standard errors in parentheses are clustered by host country in columns 1 to 3. ***, **, * denote significance level at 1%, 5% and 10% respectively.

Appendix Table A.4: Instrumenting with the Components of the Predicted Migrant Network

Estimation method Stage I dependent variable	OLS	OLS	Poisson	Negative binomial
	Migrant network			
	(1)	(2)	(3)	(4)
(log) 1960 migrant stock	0.229*** (0.052)	0.192*** (0.050)	0.229*** (0.052)	0.229*** (0.052)
$Push_{i,t} \times Pull_{i,t}$	0.025*** (0.008)	0.013* (0.008)	0.025*** (0.008)	0.025*** (0.008)
Controls and time dummies	Y	Y	Y	Y
Stage II dependent variable	(log) MA+1	(log) MA	MA	MA
Migrant network	0.253*** (0.067)	0.434*** (0.148)	0.668*** (0.118)	0.639*** (0.118)
Controls and time dummies	Y	Y	Y	Y
N	943	245	943	943
Shea partial R^2	0.23	0.26		
Underidentification test (p -value)	<0.01	<0.01		
Cragg-Donald Wald F statistic	138.6	38.45		
Kleibergen-Paap rk Wald F statistic	11.78	7.73		
Overidentification test (p -value)	0.59	0.23		

Notes: The set of control variables is described in the text. Standard errors in parentheses are clustered by host country in columns 1 to 3. ***, **, * denote significance level at 1%, 5% and 10% respectively. p -values associated with the Kleibergen-Paap rk Lagrange Multiplier statistic for the underidentification test and the Hansen J statistic for the overidentification test are reported.

Appendix Table A.5: Overseas Chinese Migrant Networks and Greenfield ODI– Poisson Model

Instrument Dependent variable	GF			Predicted migrant network
	(1)	(2)	(3)	(4)
Migrant network	0.089 (0.095)		0.090 (0.418)	-0.326 (0.311)
Chinese share of migrants		0.053 (0.489)		
Predicted migrant network				0.081 (0.077)
Controls and time dummies	Y	Y	Y	Y
Time averages of controls	Y	Y	Y	Y
Time averages of time dummies	Y	Y	Y	Y
Industry dummies	N	N	Y	N
N	966	966	8,694	943

Notes: The set of control variables is described in the text. Column 4 presents the second-stage regression results with the predicted migrant network as the instrument using the control function approach. Standard errors in parentheses are clustered by host country. ***, **, * denote significance level at 1%, 5% and 10% respectively.

Appendix Table A.6: Heterogeneity Analysis– Poisson Model

Panel A: Heterogeneity by overseas investment experience						
	(1)	(2)	(3)	(4)		
Migrant network	0.207	0.141	-0.013	0.208**		
	(0.579)	(0.571)	(0.261)	(0.089)		
First	0.865***	-0.348				
	(0.099)	(0.547)				
Migrant network × First		0.102**				
		(0.043)				
(log) Cumulative MA			-0.790	-1.150		
			(0.618)	(0.717)		
(log) Cumulative MA			1.607	3.637***		
			(1.023)	(0.738)		
Migrant network × (log) Cumulative MA				0.039		
				(0.050)		
Migrant network × (log) Cumulative MA				-0.179***		
				(0.050)		
<i>N</i>	1,932	1,932	966	966		

Panel B: Heterogeneity across firms and industries						
Subsample	Unlisted acquirer or target	Listed acquirer and target	Private-owned enterprise	State-owned enterprise	Across industries	Within industries
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant network	0.125	0.466	0.402	-0.118	0.178	0.130
	(0.640)	(0.573)	(0.570)	(0.708)	(0.926)	(0.438)
First	-0.383	0.908	1.319*	-2.312**	-0.508	-0.178
	(0.600)	(4.829)	(0.722)	(1.136)	(0.632)	(0.970)
Migrant network × First	0.108**	-0.049	-0.022	0.239**	0.128**	0.065
	(0.049)	(0.402)	(0.054)	(0.099)	(0.052)	(0.075)
<i>N</i>	1,932	1,764	1,932	1,932	1,932	1,932

Panel C: Heterogeneity across host countries						
	Regulation index		Business regulations		Bribes, favoritism	
	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant network	0.087	0.567***	-0.123	0.410	-0.049	0.386***
	(0.113)	(0.087)	(0.090)	(1.628)	(0.109)	(0.099)
First	-1.347*	0.797	-1.791**	1.177*	-1.939*	1.066*
	(0.811)	(0.653)	(0.789)	(0.685)	(1.047)	(0.581)
Migrant network × First	0.191**	0.013	0.234***	-0.018	0.243**	-0.007
	(0.084)	(0.048)	(0.073)	(0.050)	(0.108)	(0.041)
<i>N</i>	942	924	928	910	886	886

	Hiring and firing regulations		Credit market regulations		Foreign ownership restrictions	
	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant network	0.390***	-0.008	0.080	0.310	0.048	0.350***
	(0.125)	(0.099)	(0.090)	(0.710)	(0.106)	(0.101)
First	-1.121	-0.220	-0.820	0.662	-0.489	0.021
	(0.941)	(0.577)	(0.637)	(0.740)	(0.745)	(0.892)
Migrant network × First	0.182**	0.083**	0.134**	0.026	0.107*	0.077
	(0.083)	(0.041)	(0.054)	(0.056)	(0.064)	(0.073)
<i>N</i>	894	878	936	930	894	878

Notes: All regressions include controls and time dummies, time averages of controls, and time averages of time dummies. The set of control variables is described in the text. Standard errors in parentheses are clustered by host country. ***, **, * denote significance level at 1%, 5% and 10% respectively.