

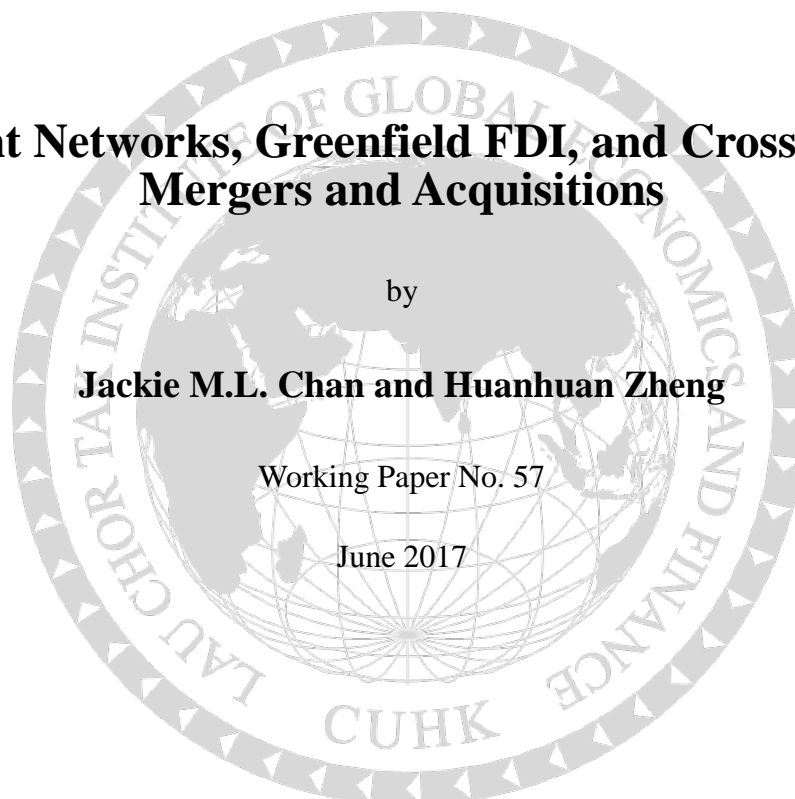
# **Migrant Networks, Greenfield FDI, and Cross-border Mergers and Acquisitions**

by

**Jackie M.L. Chan and Huanhuan Zheng**

Working Paper No. 57

June 2017



Lau Chor Tak Institute of Global Economics and Finance  
The Chinese University of Hong Kong  
13/F, Cheng Yu Tung Building, 12 Chak Cheung Street, Shatin, Hong Kong

# Acknowledgements

---

The Lau Chor Tak Institute of Global Economics and Finance is grateful to the following individuals and organizations for their generous donations and sponsorship (in alphabetical order):

## Donors

Johnson Cha	Agile Group Holdings Limited
Vincent H.C. Cheng	Asia Financial Holdings Ltd
Jonathan K.S. Choi	Bank of China (Hong Kong) Limited
Fred Hu Zulu	BCT Financial Limited
Lau Chor Tak and Lau Chan So Har	China Concept Consulting Ltd
Lawrence J. Lau	First Eastern Investment Group
Chien Lee	Hang Lung Properties Limited
Antony Leung	Henderson Land Development Co. Ltd.
Wei Bo Li	Hong Kong Exchanges and Clearing Limited
Francis Lui	Industrial and Commercial Bank of China (Asia) Limited
Robert Ng	Lai Sun Development Co., Ltd.
Simon Suen	Lau Chor Tak Foundation Limited
Allan Zeman	Sing Tao News Corporation Ltd.
	Sun Hung Kai Properties Ltd.
	Tai Sang Bank Limited
	The Bank of East Asia, Limited
	The Hongkong and Shanghai Banking Corporation Limited
	The Lanson Foundation
	Wing Lung Bank Limited

## Programme Supporters

C.K. Chow	Bangkok Bank Public Co Ltd
Alvin Chua	Bank of China (Hong Kong) Limited
Fang Fang	Bank of China Limited - Phnom Penh Branch
Eddy Fong	Bei Shan Tang Foundation
Victor K. Fung	China Development Bank
Wei Bo Li	China Soft Capital
K.L. Wong	HOPU Investment Management Co Ltd
	Industrial and Commercial Bank of China - Phnom Penh Branch
	King Link Holding Limited
	Sun Wah Group
	The Santander-K Foundation
	UnionPay International

---

# Migrant Networks, Greenfield FDI, and Cross-border Mergers and Acquisitions<sup>§</sup>

Jackie M.L. Chan<sup>\*</sup> and Huanhuan Zheng<sup>\*\*</sup>

June 2017

**Abstract:** We study the relationship between international migration and foreign direct investment (FDI), differentiating between the two modes of entry: greenfield FDI and cross-border mergers and acquisitions (M\&A). Each mode may face different formal and informal information barriers, from regulations, laws, cultural differences, as well as the information intensity required in the transaction. Using data on overseas Chinese migrant networks and outward FDI from 2003-2014 to over 100 countries, we estimate a gravity equation for FDI. We find a positive association between overseas Chinese migrant networks and Chinese outward FDI, and importantly, they are more closely associated with M\&A than greenfield FDI. The relationships are more pronounced where formal and informal barriers are stronger: in host countries with tighter regulations and those that are more culturally distant, in service industries and cross-industry FDI, and for investment from privately held companies and privately-owned enterprises. The results are robust to various specifications and estimation methods, including an instrumental variables approach that addresses potential endogeneity concerns.

**JEL:** F21, F22, F23

**Keywords:** Chinese migrant networks, greenfield, mergers and acquisitions, foreign direct investment, information

---

<sup>§</sup> © 2017 Lau Chor Tak Institute of Global Economics and Finance, The Chinese University of Hong Kong

<sup>\*</sup> Department of Economics, 9/F, Esther Lee Building, Chinese University of Hong Kong, Shatin, N.T., Hong Kong, jmlchan@cuhk.edu.hk. An earlier version of this paper was titled “Migrant networks and FDI: Evidence from Greenfield Investment and M\&A”. The authors are grateful for the suggestions provided from Baptist University and participants of the CUHK Conference on Trade and Development.

<sup>\*\*</sup> Department of Economics, Lau Chor Tak Institute of Global Economics and Finance, Chinese University of Hong Kong, Shatin, N.T., Hong Kong, arwenzh@gmail.com. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the Institute.

# 1 Introduction

It has long been recognized that ethnic and international migrant networks, with their familiarity in language, culture, and regulations across countries, can help overcome barriers that accompany international transactions (e.g., Rauch, 2001; Gould, 1994). Existing studies document the positive role of ethnic or migrant networks in promoting international trade (e.g., Rauch and Trindade, 2002), as well as foreign direct investment (FDI) (e.g., Javorcik et al., 2012).<sup>1</sup> However, for the latter, heterogeneity has largely been overlooked, with the analysis made mostly at the aggregate level. In this paper, we examine the effects of international migration on FDI, and in particular, study the impact of overseas Chinese migrant networks on outward FDI from China, differentiating between the two modes of entry: greenfield and brownfield FDI.

Greenfield FDI refers to cross-border investment in which a parent company establishes new production facilities in the host country, whereas brownfield FDI is typically associated with cross-border mergers and acquisitions (M&A), in which a firm acquires an existing company overseas.<sup>2</sup> Figure 1 plots the value and frequency of investment by mode of entry, comparing between developed and developing countries.<sup>3</sup> As Panel (a) indicates, for source countries, greenfield investment dominates brownfield investment in values and counts, regardless of the country’s income level. The pattern is not as uniform for host countries (Panel (b)). Moreover, the majority of FDI (roughly two-thirds by value and even higher by counts), originates from developed economies.

Even though both forms of FDI are conducted to gain access to a foreign market, the entry barriers that greenfield FDI and cross-border M&A face can be quite different. While greenfield FDI can be viewed as a firm’s expansion across national boundaries, a M&A instead crosses organization boundaries. Thus, along with the monetary costs of setting up a subsidiary, a firm pursuing greenfield investment must obtain knowledge of foreign management and navigate country differences in laws, regulations, and culture. However, aside from the setup cost, the differences in business environment also exist for brownfield FDI, even though local expertise within the firm may help resolve some of the problems. Since M&A involve the transfer of ownership, there are also many problems associated with negotiation, bargaining, and contract enforcement that are not present in greenfield investment. In addition to costs for 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 subsidiary

---

<sup>1</sup>The UN Conference on Trade and Development (UNCTAD) defines foreign direct investment as an “investment made to acquire lasting interest in enterprises operating outside of the economy of the investor”, and requires that the “single foreign investor either owns 10 per cent or more of the ordinary shares or voting power of an enterprise ... or owns less than 10 per cent of the ordinary shares or voting power of an enterprise, yet still maintains an effective voice in management.”

<sup>2</sup>We use the terms brownfield FDI and cross-border M&A interchangeably in this paper.

<sup>3</sup>The data are retrieved from Annex Tables 9, 10, 11, 12, 18, 19, 21, and 22 of (UNCTAD, 2016). The classification of developed and developing economies here follows the UN dataset. The patterns are similar if the average between 2006 and 2010 are computed instead. Similar data patterns are documented in, for instance, Davies et al. (2015).

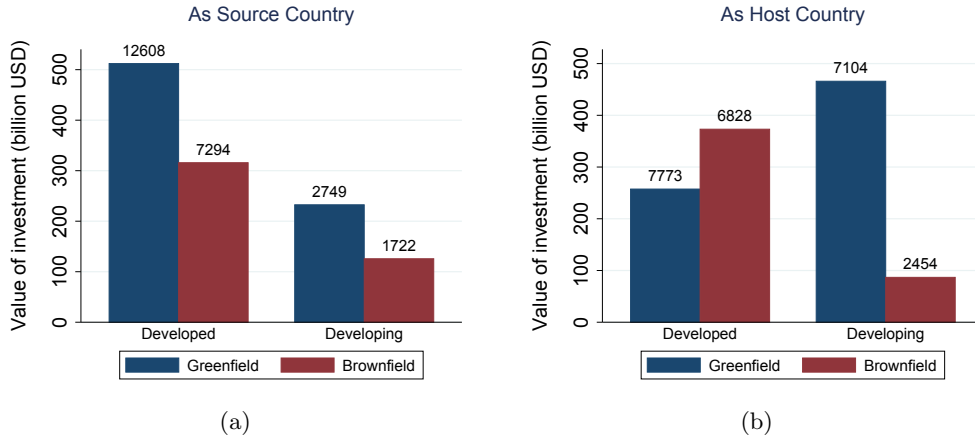


Figure 1: The value of greenfield and brownfield (i.e., M&A) foreign direct investment (in billions of USD) averaged between 2011 and 2015 for developed and developing economies. The average number of investment projects is displayed above each bar. Panel (a) presents the statistics as a source or purchaser country, while Panel (b) presents the statistics as a host or seller country. (Source: UNCTAD World Investment Report.)

is established or acquisition 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 demand is key for marketing.

Furthermore, the mode of entry for foreign investment can have significantly different implications for the host country. While greenfield FDI creates job opportunities and technology spillovers at new production facilities, brownfield FDI tends to be associated with little job creation and absorbs technology from acquired companies. Hence, host countries generally welcome greenfield FDI more than brownfield FDI (UNCTAD, 2014). Governments may also restrict foreign acquisitions due to concerns of national security, job loss, and the protection of firms in strategic industries (Bertrand et al., 2012; UNCTAD, 2016).

This paper empirically studies the relationship between overseas migrant networks and outward FDI (OFDI) to the destination countries, and importantly, how it varies by investment entry mode. We focus on the recent growth of Chinese migration and investment abroad, and provide novel evidence using data on overseas Chinese migrant stocks and deal-level Chinese greenfield and brownfield FDI to 135 countries between 2003 and 2014.<sup>4</sup> To the best of our knowledge, this is the first paper to individually examine how migrant networks influence greenfield and brownfield cross-border investment. With a gravity model, we find that the size of the Chinese migrant network overseas has a large and significant impact on Chinese investment into the host country, for both greenfield FDI and cross-border M&A. Importantly, migrants appear to be more important in alleviating the problems associated with the latter. These findings remain robust when alternative estimation methods are used. In particular, we employ

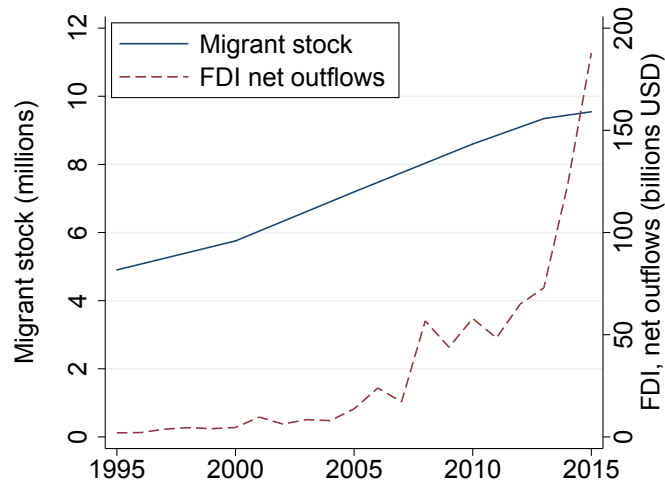
<sup>4</sup>Data on greenfield FDI and M&A are obtained from fDi Intelligence and Thomson-Reuters SDC Platinum, respectively; see Section 3 for details. SDC Platinum is an older database that has been utilized in many papers (e.g. di Giovanni, 2005; Head and Ries, 2008), and the same greenfield FDI data is analyzed by the UN in their annual report on world investment (UNCTAD, 2016).

an instrumental variables (IV) strategy address endogeneity concerns, where our instruments include the historical migrant network from 40 years ago, and countries' contemporaneous policy on dual citizenship, immigration strictness, and residents' hostility towards immigrants. We also construct an instrument based on Burchardi et al. (2016). Using global bilateral migration flows, migrant shares are predicted from push factors that do not depend on the host-country, and pull factors that are independent of China. Regressions from the IV strategy strongly reinforce our baseline results.

To understand why migrant networks may differ in promoting greenfield and brownfield FDI, we investigate their role in alleviating informal and formal barriers associated with international transactions. Using the detailed disaggregate deal-level FDI data, we find substantial heterogeneity across firms, industries, and various host-country dimensions. At the industry level, we find a more pronounced effect for investment in service industries, which are typically more knowledge intensive (Coff, 1999), for FDI towards a different industry from the investing company. A stronger effect is also observed in host countries that are more culturally distant from China with tighter regulations, and again for M&A than greenfield investment. At the firm-level, migrant networks promote more FDI from or between companies that are privately held (i.e., not publicly traded), as well as from privately-owned enterprises. Taken together, the empirical evidence strongly suggests that overseas migrants and their networks provide knowledge regarding firms, industries, and countries to facilitate international investment, and supports anecdotal evidence that for brownfield investment may be more information-intensive than greenfield investment.

For example, the effect of migrant networks on OFDI is more pronounced for investment from or between companies that are privately held (i.e., not publicly traded) with more information withheld from the public. It is also larger for privately-owned enterprises, which utilize overseas Chinese networks more actively than state-owned enterprises (CCPIT, 2015). In addition, at the industry level, the positive relationship is stronger in tertiary industries, which are typically more knowledge intensive and face greater foreign ownership restrictions than the secondary or primary industries (Coff, 1999; World Bank, 2010), and for FDI towards a different industry from the investing company. Lastly, cross-border investment increases with the size of migrant networks to a greater extent in host countries that are more culturally distant from China with tighter regulations. Again, we observe stronger effects on M&A than on greenfield investment. Taken together, the empirical evidence presented strongly suggests that overseas migrants and their networks provide information regarding firms, industries, and countries to facilitate international investment, and is consistent with the anecdotal evidence and idea that brownfield investment face greater information barriers than greenfield FDI.

The focus on Chinese migration and investment is motivated by the observation that along with exceptional growth, China has also seen large movements of labor and capital abroad. As Figure 2 shows, the international Chinese emigrant stock roughly doubled between 1995 and 2015, from 4.9 to 9.5 million. Furthermore, China has become a huge investor abroad, as net FDI outflows have risen to historic levels, and it has become the third largest investing



(a)

Figure 2: Chinese migrant stock abroad and FDI, net outflows. (Sources: UN Global Migration Database and the World Bank World Development Indicators.)

home economy behind the US and Japan (UNCTAD, 2016). The values of both greenfield and brownfield FDI have also doubled in the last decade. The research question is also particularly relevant for a developing country like China, as Chinese firms may have less experience operating abroad, and barriers such as language differences are especially large.

The role of ethnic and migrant networks in overcoming informal barriers and facilitating economic activity across borders has been studied in various contexts, most notably in international trade (e.g., Gould, 1994; Head and Ries, 1998; Combes et al., 2005; Dunlevy, 2006; Parsons and Vézina (forthcoming)). Focusing on overseas Chinese ethnic networks, Rauch and Trindade (2002) show heterogeneous effects depending on the goods being traded (i.e., differentiated commodities homogeneous). Our work contributes most directly to the literature on migration and FDI. Much of it has focused on immigration into the US. For example, Javorcik et al. (2012) find a positive relationship between immigrant shares in the US and FDI towards the migrants' origin countries for 1990 and 2000, especially for more educated immigrants.<sup>5</sup> Importantly, they address the endogeneity issue by using an instrumental variables approach.<sup>6</sup> More recently, Burchardi et al. (2016) and Cohen et al. (2016) utilize the historical geographic distribution of ethnic groups in the US to address endogeneity. In particular, using a reduced-form model of migration based on push and pull forces, Burchardi et al. (2016) construct an instrument by relying on the ancestry composition of migrants.

There have been few studies that have examined our research question for other countries like China, despite the large numbers of emigrants coming from these countries.<sup>7</sup> Most importantly,

<sup>5</sup>Earlier work by Bhattacharya and Groznik (2008) also found US outward FDI to a particular destination to be correlated with the size of the immigrant group from the same country. Kugler and Rapoport (2007) document a similar pattern over time, but also show a negative contemporaneous correlation.

<sup>6</sup>Their instruments for immigrant entry into the US are the cost of obtaining a passport in the origin country, the historical share of migration, distance to the EU, the presence of a US military base, and dual citizenship.

<sup>7</sup>Using the ethnic Chinese population data in 1990 from Rauch and Trindade (2002), Gao (2003) and Tong

we contribute to this area of research by using deal-level FDI data to differentiate between greenfield and brownfield FDI, and provide new evidence on the role of migrant networks for each type of investment. Our results on the heterogeneous effects across firms, industries, and host countries also explicitly test and reveal how migrants might promote FDI.

Our paper also complements existing empirical evidence on the choice of foreign market entry mode (e.g., Hennart and Park, 1993; Nocke and Yeaple, 2008), as well as theoretical predictions on how this choice is affected by changes in FDI policy and trade liberalization (Qiu and Wang, 2011; Stepanok, 2015). Using the same greenfield FDI dataset and a related database on M&A, Davies et al. (2015) also find M&A are more affected by cultural barriers.<sup>8</sup> More broadly, our paper informs studies that examine the determinants of FDI. Our empirical analysis controls for many of the various factors which have been empirically identified to be correlated with multinational activity (e.g., Blonigen and Piger, 2014).<sup>9</sup> Lastly, our work adds to the recent growing empirical research on Chinese OFDI.<sup>10</sup> Using data on FDI from Zhejiang province, Chen et al. (2016) find privately-owned multinationals to be discriminated in the domestic Chinese market and thus engage in investment and production abroad.

The rest of paper is organized as follows. In Section 2, we discuss the entry barriers of greenfield and brownfield FDI conceptually. Section 3, we describe the data sources and present some descriptive statistics. Section 4 outlines the empirical framework and presents the empirical results, including robustness checks and instrumental variables estimates. Lastly, Section 5 concludes.

## 2 Entry barriers of greenfield and brownfield FDI

## 3 Data

### 3.1 FDI data

We draw data from a variety of sources. First, transaction-level data on Chinese outward greenfield FDI and cross-border M&As are obtained from Financial Times Ltd. fDi Intelligence

---

(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. They find that ethnic Chinese FDI firms under-perform non-ethnic Chinese FDI firms. Buch et al. (2006) study German inward FDI, and find more FDI into states with a large foreign population from the same origin country. Recent work by Burchardi and Hassan (2013) shows how social ties increase economic development and investment by examining links between West and East Germany after the Berlin Wall fell.

<sup>8</sup>For M&A, Davies et al. (2015) rely on Zephyr, a database from another business information publisher, Bureau van Dijk. Using global FDI data, they find M&A are more sensitive to geographic and cultural barriers and short-run variations like a currency crisis, while greenfield FDI is affected more by long-run factors such as institutional development.

<sup>9</sup>Other papers in this large literature include Bevan and Estrin (2004); Portes and Rey (2005); Bénassy-Quéré et al. (2007); Head and Ries (2008); Chang (2014); Lee (2016).

<sup>10</sup>For example, Buckley et al. (2011), Kolstad and Wiig (2012), and Huang and Wang (2011, 2013) find that Chinese OFDI is asset-driven and attracted towards big markets with natural resources and advanced technology. As well, firm productivity has been shown to increase the likelihood and size of investment overseas (e.g., Chen and Tang, 2014; Wang et al., 2016; Tian and Yu, 2015).



and Thomson-Reuters Security Data Company (SDC) Platinum, respectively.<sup>11</sup> Both datasets are commonly used to analyze cross-border greenfield and brownfield investment. For example, UNCTAD (2016) rely on fDi Intelligence for their annual report on world investment, and the M&A data in SDC Platinum has been examined in numerous papers (e.g., Rossi and Volpin, 2004; di Giovanni, 2005; Head and Ries, 2008).<sup>12</sup> We limit our sample to the years from 2003 to 2014, because 2003 is the earliest year that fDi Intelligence recorded greenfield investments. Overall, our sample includes 3184 greenfield investment projects across 136 host countries (with an average of 265 projects and total valuation of 26 billion USD annually), and 1393 cross-border M&A deals in 90 countries (with an average of 116 deals and reported transaction value of 18 billion USD per year).<sup>13</sup>

The distribution of greenfield FDI projects and M&A deals across host countries is listed in Appendix Table 1 and displayed as histograms in Appendix Figure 1. The number of greenfield projects for any given host country is typically larger than the number of cross-border M&A deals, with notable exceptions being Australia and Canada. Between 2003 and 2014, over 60% of host countries receive less than a total of 10 greenfield FDI investment projects from China, and close to 80% see less than 10 M&A with Chinese companies.<sup>14</sup> There are, however, some host countries that China invests heavily in. Specifically, the five most popular destinations for Chinese greenfield OFDI over this decade are Germany, US, UK, Hong Kong, and India; likewise, for brownfield, they are: Hong Kong, US, Australia, Canada, and Germany.

Next, Figure 3 displays the time-series variation in the number and total value of greenfield and brownfield investments. Over this decade, both greenfield FDI and M&A exhibit a rising trend, especially after the 2007 global financial crisis. Panel (a) shows that the total number of Chinese greenfield investment projects is consistently around two to three times larger than the number of cross-border M&A deals. In terms of total transaction value as plotted in Panel B, brownfield OFDI sometimes exceeds that of greenfield OFDI, which implies that the size of brownfield projects is, on average, greater than greenfield projects.

Due to the large state-sector in the Chinese economy, the decomposition between state-owned enterprises (SOE) and privately-owned enterprises (POE) is interesting to examine. In Section 4.5.3, this dimension will also be studied to understand the effect of migrants for different types of firms. For each investing (acquiring) company, we collect its ownership structure from Orbis, a database managed by Bureau van Dijk which contains information on 200 million

---

<sup>11</sup>Specifically, fDi Intelligence provides the number and value of announced greenfield FDI projects, while SDC Platinum records both completed cross-border M&A deals.

<sup>12</sup>UNCTAD (2016) 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 (2016) are highly correlated at 0.91 and 0.85, respectively. SDC Platinum captures 38 to 80% of the number of M&A deals in UNCTAD (2016), and 40 to 276% of the value.

<sup>13</sup>For both the greenfield FDI and M&A data, the value of transactions is not reported in some instances due to confidentiality. Thus, while transaction counts can be always computed, deal valuations are not available for all transactions. In the case of greenfield FDI, the data provider, fDi Intelligence, estimates the value of the deal if it is not disclosed by the investing company. However, estimates for the values of M&A investments are not included.

<sup>14</sup>These statistics are computed using the sample of host countries with strictly positive greenfield or brownfield FDI. There are 81 countries with both greenfield and brownfield Chinese investment activity, 55 countries with only greenfield investments and 9 countries with only brownfield investments.

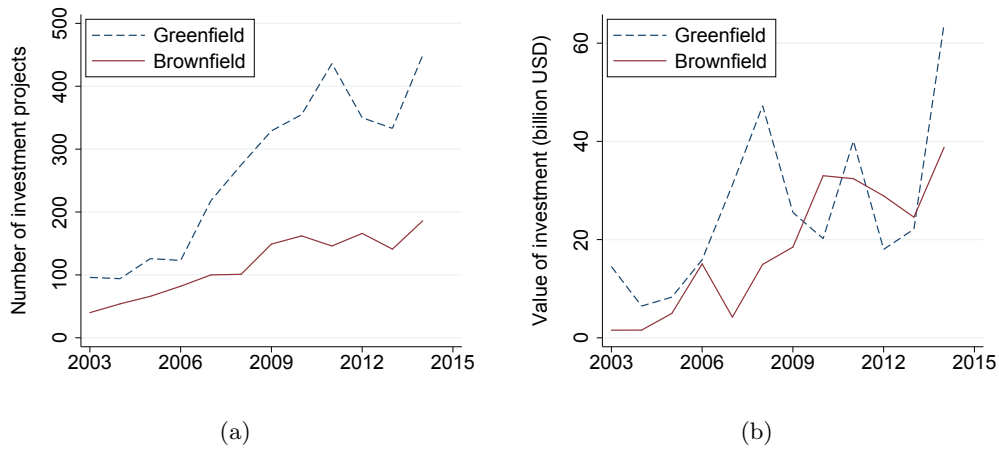


Figure 3: Panels (a) and (b) plot the time-series of the number and value of Chinese greenfield and brownfield investment projects, respectively.

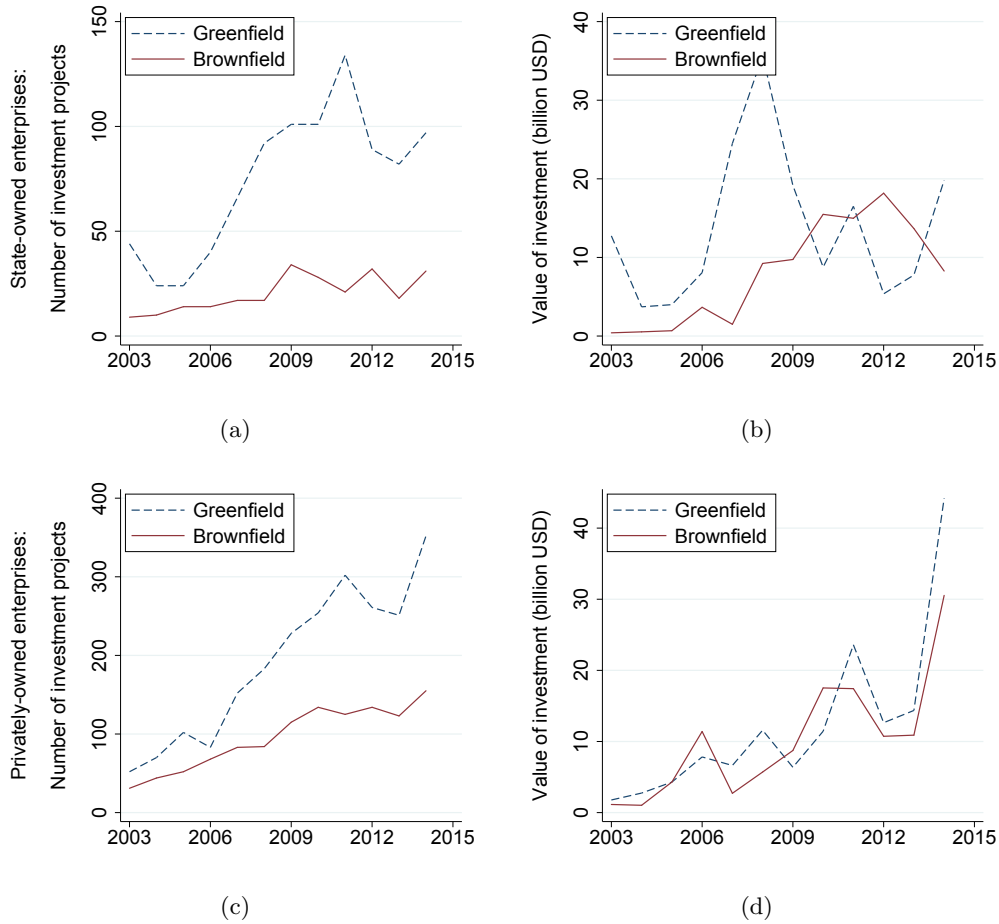


Figure 4: Panels (a) and (b) plot the time-series of the number and value of Chinese greenfield and brownfield investment projects from state-owned enterprises, respectively. Panels (c) and (d) plot the time-series of the number and value of Chinese greenfield and brownfield investment projects from privately-owned enterprises, respectively.

companies worldwide. A firm is classified as a state-owned enterprise (SOE) if at least 25.01% of its ownership belongs ultimately to the Chinese government, and a privately-owned enterprise (POE) otherwise.<sup>15</sup> Decomposing OFDI by the ownership structure of Chinese investors, we observe in Figure 4 a similar pattern that there is a larger number of greenfield projects of smaller size relative to brownfield investments, regardless of whether the OFDI originates from SOEs or POEs. Moreover, while there tends to be more investment projects from POEs, the value of their projects is generally smaller than those from SOEs.

In the empirical analysis, we also employ a measure of aggregate cross-border investment flows from UNCTAD Bilateral FDI Statistics. The sample contains FDI outflows from China to over 200 host economies between 2003 and 2012. These three complementary datasets of FDI will allow for separate examinations of aggregate and disaggregate FDI by mode of entry.

### 3.2 Bilateral migrant stock data

Data on bilateral migration are obtained from the Global Migration Database provided by the UN Department of Economic and Social Affairs. Statistics on bilateral migrant stocks are available every five years from 1990 to 2015, with an additional year in 2013. Our key explanatory variable is the share of Chinese migrants in the host country, i.e., the number of Chinese emigrants to host country  $j$  in a particular year normalized by the host country's population. Data on the host country's population is retrieved from the World Bank World Development Indicators (WDI). For years in which the migration data are not available, we carry the value of migrant share forward until the new data becomes available. Because of the small within-country time variation, the regression estimates are almost identical if we interpolate linearly the values for the years in which the data is unavailable. Thus, for example, the migrant share in 2006 to 2009 will take the same value as 2005. Since there are no sudden surges in emigration to a particular destination, Chinese migrant shares are relatively stable over time. The temporal variation of migrant shares is very low. The simple average (across all countries) of the standard deviation of migrant shares across 12 years is around 0.06%. For comparison, the mean migrant share across all countries and years is close to 0.9%.<sup>16</sup>

### 3.3 Other data

In the regression analysis, we control for host country characteristics that may influence FDI to mitigate the concern of omitted variable bias. These include traditional gravity variables like

---

<sup>15</sup>Our results in Section 4.5.3 are robust to changing the equity threshold of the global ultimate owner to 50.01%.

<sup>16</sup>Gao (2003) and Tong (2005) examine the relationship between Chinese *ethnic* networks abroad and aggregate FDI using data on ethnic Chinese populations in 1990. A comparison can be made between this dataset and the UN database for the overlapping year of 1990. As Appendix Figure 2 illustrates, the correlation between the two variables is very high (at 0.684). One would expect, if the data is collected accurately, that the ethnic populations are larger than the number of migrants. This is true for 36 out of 49 observations. For observations below the 45 degree, this suggests either the ethnic population was underestimated or the migrant population was overestimated.

market size (as proxied by GDP) and geographic distance.<sup>17</sup> To account for the market-seeking and growth-seeking motives of Chinese OFDI (e.g., Deng, 2004; Buckley et al., 2011), we control for income (as measured by GDP per capita) and real GDP growth in the host country; both variables are drawn from the WDI. To control for cultural similarity that makes cross-border investments more likely (e.g., Dunlevy, 2006; Davies et al., 2015), we add a dummy for common language that equals to 1 if at least 9% of the population in the host country speaks the same language with China, and a dummy for a common legal system origin. The former is obtained from CEPII (along with distance), and the latter from La Porta et al. (1999).

Furthermore, financial development is included since it can potentially facilitate the international expansion of firms through FDI (e.g., Desbordes and Wei, 2014); it is measured by private credit from deposit money banks and other financial institutions divided by GDP (Beck et al., 2000). We follow Rossi and Volpin (2004) and Bekaert et al. (2004), among others, and use the sum of indices for corruption, law and order, and bureaucratic quality from the International Country Risk Guide (ICRG) as a proxy for institutional quality. Lastly, following Blonigen et al. (2007), we use trade openness, the ratio of exports plus imports to GDP (from WDI), to control for the degree of business interactions with the rest of the world. The definition and sources of variables used in this paper are listed in Appendix Table 2, and summary statistics are provided in Appendix Table 3,

## 4 Empirical framework and results

### 4.1 Empirical framework

Following the literature on international trade and the determinants of FDI, we utilize a simple gravity model for bilateral FDI. The gravity equation has been reasonably successful in fitting the observed data of cross-country trade and FDI flows (e.g., Bénassy-Quéré et al., 2007; Blonigen and Piger, 2014). The gravity equation can be specified as follows:

$$FDI_{ij} = b_0 GDP_i^{\beta_1} GDP_j^{\beta_2} Dist_{ij}^{\beta_3} \exp(\beta_{\mathbf{X}} \mathbf{X}_{ij} + \varepsilon_{ij}),$$

where  $FDI_{ij}$  is investment from country  $i$  to country  $j$ ,  $GDP$  captures market size, and  $Dist_{ij}$  is the geographic distance between the two countries. To mitigate concerns of omitted variable bias,  $\mathbf{X}_{ij}$  is a vector that includes other potential determinants of cross-border investment, which are either bilateral or host/home-country specific. Lastly,  $\varepsilon_{ij}$  is the error term. For each year, the gravity equation is log-transformed:

$$\log(FDI_{ijt}) = \beta_0 + \beta_1 \log(GDP_{it}) + \beta_2 \log(GDP_{jt}) + \beta_3 \log(Dist) + \beta_{\mathbf{X}} \mathbf{X}_{ijt} + \varepsilon_{ijt}.$$

To examine the heterogeneous effects at the aggregate and disaggregate level by mode of entry, our dependent variable includes the number and value of greenfield investment projects

---

<sup>17</sup>An alternative measure of market size is population. In unreported regressions, we confirm our results are robust to this measure. These results are available upon request.

( $\#GF$ ,  $GF$ ), the number and value of cross-border M&A deals ( $\#BF$ ,  $BF$ ) from China to the host country, and aggregate FDI outflows. Since China only has positive investment into a subset of host countries every year, we apply a very standard approach to deal with zero FDI flows. In our baseline specification, the dependent variable  $\log(FDI_{ijt})$  is replaced with  $\log(FDI_{ijt} + 1)$ , i.e., a value of zero FDI projects would be replaced by one FDI project, and a value of zero FDI flows becomes \$10,000 USD, which is the smallest value of FDI observed in the data. In robustness checks, we demonstrate our findings are robust to alternative estimation methods, including OLS without replacing zeros with ones, as well as the Poisson pseudo-maximum-likelihood (PPML) estimator. Furthermore, to mitigate the concern of reverse causality, all the time-varying explanatory variables are lagged by one year (e.g., Chang, 2014). In Section 4.4, we validate our results with an instrumental variables strategy to address potential endogeneity issues.

Thus, focusing on the OFDI from China, the following baseline specification is estimated using OLS to study the impact of overseas Chinese migration networks on different types of OFDI:

$$\begin{aligned} \log(FDI_{jt} + 1) = & \alpha_0 + \alpha_1 \log(GDP_{j,t-1}) + \alpha_2 \log(Dist) + \alpha_3 \text{Migrant share}_{j,t-1} \\ & + \alpha_{\mathbf{X}} \mathbf{X}_{j,t-1} + \epsilon_{jt}. \end{aligned} \quad (1)$$

Our regressor of interest is *Migrant share*, defined as the number of overseas Chinese migrants in host country  $j$  as a ratio of the total population in country  $j$ . If Chinese migrants overseas do help Chinese investors overcome informal and formal barriers and increase the amount of FDI, then  $\alpha_3$  should be positive. The vector  $\mathbf{X}_{j,t-1}$  also includes year fixed effects to absorb time-specific changes in FDI flows common to all countries. Standard errors of the estimated coefficients allow for clustering of observations by host country.

## 4.2 Baseline results

Table 1 reports the baseline regression results from estimating Eq. (1) with OLS. The regressors consist of standard gravity equation variables (i.e., (log) GDP, (log) GDP per capita, (log) Distance, Common language, and Common legal system) and year fixed effects. This subset of control variables allows countries such as Macau to be included in the estimation sample, as is done in Table 2 below for comparison. Throughout the empirical analysis, we exclude tax haven countries such as Panama, the British Virgin Islands, the Cayman Islands. Table 1 also excludes both Macau and Hong Kong, which have unique status as special administrative regions of China, while in Table 2, we compare regressions with and without these two countries. Thus, in Table 1, the sample contains 135 countries between the years 2003 and 2014 in columns 1 to 4 for disaggregate FDI, and 124 countries from 2003 to 2012 in column 5 for aggregate FDI.

Consistent with the hypothesis that cross-border migrant networks help overcome barriers associated with international transactions, we find that *Migrant share* is positively and significantly associated with different measures of OFDI: both greenfield and brownfield FDI, and

Table 1: Overseas Chinese Migrant Networks and Chinese Outward FDI

Dependent variable	log (FDI + 1)				
	# GF (1)	GF (2)	# BF (3)	BF (4)	UN FDI (5)
Migrant share	14.590*** (4.056)	59.781*** (19.787)	24.703*** (7.679)	123.696*** (34.557)	82.551*** (29.578)
(log) GDP	0.231*** (0.024)	1.233*** (0.084)	0.113*** (0.021)	0.569*** (0.082)	0.904*** (0.096)
(log) GDP per capita	-0.067*** (0.025)	-0.526*** (0.128)	0.005 (0.014)	0.048 (0.081)	-0.810*** (0.163)
(log) Distance	-0.163** (0.066)	-1.078*** (0.349)	-0.052 (0.052)	-0.242 (0.329)	-1.294*** (0.416)
Language	0.230 (0.225)	2.189* (1.149)	-0.167 (0.336)	-0.381 (1.368)	-0.130 (1.175)
Legal system	-0.038 (0.095)	0.407 (0.500)	-0.111** (0.047)	-0.650** (0.287)	-0.231 (0.629)
Year fixed effects	Y	Y	Y	Y	Y
Number of countries	135	135	135	135	124
Observations	1,483	1,483	1,483	1,483	1,240
R-squared	0.511	0.455	0.387	0.298	0.391

*Notes:* The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), likewise for M&A ( $\#BF$  and  $BF$ ), and \$10,000 plus the value of aggregate FDI ( $UN FDI$ ). All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.

both for counts and values. This is the first key result, which we confirm below with alternative estimation methods and an IV strategy. The magnitudes of the coefficients are also large and economically meaningful. All else equal, a one-standard deviation increase (0.73%) in the share of Chinese migrants in the host country's population raises the count measures  $\#GF$  and  $\#BF$  by roughly 10.7% and 18.0%, respectively, and increases the investments value measures  $GF$  and  $BF$  by 43.6%, 90.3%, respectively.

Table 1 also shows that the positive relation between migrants networks and OFDI exhibits substantial heterogeneity depending on the mode of entry. We hypothesize that the positive migrant effect is stronger for brownfield FDI compared to greenfield FDI because of the higher barriers encountered for cross-border M&A. This second key result is indeed confirmed in Table 1. Specifically, the coefficient of *Migrant share* in the regression of column 3 (4) is larger than that in column 1 (2). The difference of the marginal effects is statistically significant at 5% level with a  $\chi^2$ -statistic (and associated p-value) of 4.17 (0.04) for values, while the difference for counts is less precisely estimated with a  $\chi^2$ -statistic of 2.66 (0.10). However, both  $\chi^2$ -statistics are substantially higher when more control variables are included in Table 4 below. Thus, the results generally suggest that overseas Chinese stocks are more closely related to brownfield OFDI than to greenfield OFDI. Intuitively, cross-border M&As require intensive knowledge and information of the capabilities, preferences, and potential synergies of the acquiring and target firms that are from different cultural backgrounds, therefore, benefit more from the presence of overseas Chinese networks that bridge the gap.

Lastly, in Table 1 column 5, we rely on aggregate FDI statistics obtained from the UNCTAD

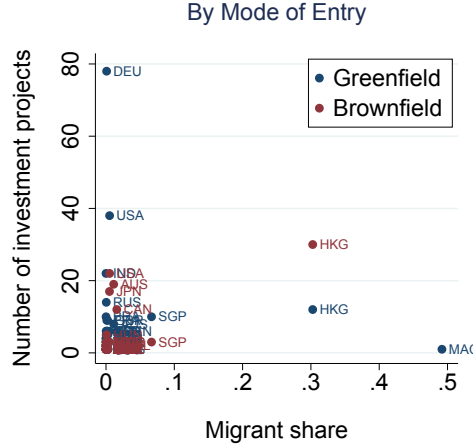


Figure 5: Number of Chinese greenfield and brownfield investment projects in 2010.

for the dependent variable. Due to data availability, the sample is slightly shorter from 2003 to 2012. Because both greenfield and brownfield FDI are measured as flow variables, we focus on aggregate FDI flows. It is clear that the positive relationship between FDI and outward migration continues to hold at the aggregate level. We also show this result remains with the stock of FDI (from the same UNCTAD database) as dependent variable instead of flows; this is available upon request.

Next, we examine how the results vary depending on the selection of countries into our sample. Figure 5 reveals two countries, Macau and Hong Kong, have very high migrant shares (above 30%) compared with the rest of the sample. These potential outliers were excluded in our baseline regression. Moreover, these two special administrative regions are contiguous to Mainland China, and while they possess their own independent legal system, the companies in these cities are often conduits for Chinese business strategies. Therefore, their FDI inflows from China may be motivated by different reasons. Thus, in Table 2, we compare the results with and without Macau and/or Hong Kong. All regressions include the same set of control variables, but for space considerations, we present the coefficient estimates of our key explanatory variable, *Migrant share*. In Panel A, the inclusion of both cities essentially removes the positive association between migrant networks and greenfield or brownfield FDI. Examining Panels B and C, it is clear that the insignificance is due to the presence of Macau. The positive relationship generally holds with Hong Kong included and Macau excluded, with statistical significance at the 1% level in 3 out of 4 columns. With the sample in Panel C, a one-standard deviation (2.79%) increase in Chinese migrant shares raises greenfield and brownfield FDI counts (values) by roughly 9.2% and 24.4% (8.2% and 75.4%), respectively.<sup>18</sup> This sharp contrast between the inclusion of Macau and Hong Kong is not surprising given the lack of FDI towards Macau as exhibited in Figure 5. In fact, in 2010, there were zero FDI projects invested in Macau by

<sup>18</sup>While no direct comparison exists in the literature on emigration and outward FDI *flows*, the results from Gao (2003) are perhaps the most relatable. He finds that a one percentage point increase in Chinese *ethnicity* shares, with Hong Kong as a FDI source country included, is associated with a 6% increase in Chinese *inward* FDI. For both Javorcik et al. (2012) and Tong (2005), the dependent variable is FDI *stock*.

Table 2: Sample Selection with and without Macau and Hong Kong

Panel A: With Macau and Hong Kong					
	# GF	GF	# BF	BF	UN FDI
	(1)	(2)	(3)	(4)	(5)
Migrant share	-0.798 (1.311)	-4.660* (2.683)	0.365 (2.667)	0.388 (9.588)	5.379 (4.429)
Observations	1,507	1,507	1,507	1,507	1,507
R-squared	0.508	0.454	0.360	0.282	0.452
Panel B: With Macau and without Hong Kong					
	# GF	GF	# BF	BF	UN FDI
	(1)	(2)	(3)	(4)	(5)
Migrant share	-1.462** (0.580)	-5.412*** (2.051)	-1.036 (1.019)	-3.532 (5.696)	3.752 (2.687)
Observations	1,495	1,495	1,495	1,495	1,495
R-squared	0.497	0.445	0.316	0.247	0.437
Panel C: With Hong Kong and without Macau					
	# GF	GF	# BF	BF	UN FDI
	(1)	(2)	(3)	(4)	(5)
Migrant share	3.284*** (0.728)	2.925 (3.676)	8.735*** (1.111)	27.019*** (7.559)	17.728*** (3.880)
Observations	1,495	1,495	1,495	1,495	1,495
R-squared	0.524	0.459	0.476	0.322	0.455

*Notes:* The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), likewise for M&A ( $\#BF$  and  $BF$ ), and \$10,000 plus the value of aggregate FDI ( $UN FDI$ ). All regressions include the set of gravity equation control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.

Chinese companies, and only one single M&A.

### 4.3 Robustness

We perform a series of additional tests that demonstrate our main results are robust to alternative specifications. First, in Table 3, we estimate our baseline regression equation using two different methods. Again, all regressions include a set of control variables, but are not reported for the sake of space. In Panel A, we use simple OLS, without replacing zeros in the dependent variable with ones. The sample size is reduced considerably as a result, but the results are similar both qualitatively as well as quantitatively. Moreover, in Panel B, we deal with zero or missing FDI flows in the data using another standard method. Specifically, we apply the Poisson pseudo-maximum-likelihood (PPML) estimator from Santos Silva and Tenreyro (2006), which simultaneously deals with the problem of zeros in the dependent variable and is consistent in the presence of heteroskedasticity. With the exception of the number of greenfield projects in Panel B column 1, which is positive but not precisely estimated, the association between the size of the migrant network and FDI remains statistically significant. Thus, even though the magnitudes of the coefficients differ from the OLS estimates, the general findings are consistent. There is a positive relationship between migration and OFDI from China, and it is stronger for



Table 3: Robustness Checks with Alternative Estimation Methods

Dependent variable	Panel A: OLS			
	log (FDI)			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	15.531*** (3.045)	39.082*** (12.056)	29.887*** (10.429)	61.504*** (11.341)
Observations	554	554	287	202
R-squared	0.504	0.237	0.352	0.288

Dependent variable	Panel B: Poisson pseudo-maximum-likelihood (PPML)			
	FDI			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	11.286 (8.716)	18.173** (7.824)	76.713** (37.970)	100.989*** (22.061)
Observations	1,483	1,483	1,483	1,483
R-squared	0.527	0.324	0.693	0.335

Dependent variable	Panel C: Estimated destination fixed effects			
	log (FDI + 1)			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	7.276 (5.018)	17.131 (24.003)	29.603*** (10.133)	147.095*** (43.097)
Observations	1,376	1,376	1,376	1,376
R-squared	0.545	0.479	0.406	0.310

*Notes:* The dependent variables are Chinese outward FDI measured by (log) number and value of greenfield investment projects ( $\#GF$  and  $GF$ ) and likewise for M&A ( $\#BF$  and  $BF$ ) in Panel A, number and value of greenfield investment projects and likewise for M&A in Panel B, and (log) one plus the number and \$10,000 plus the value of greenfield investment projects and likewise for M&A in Panel C. All regressions include the set of gravity equation control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.

brownfield investment as opposed to greenfield investment.

Next, to show our results are not driven by unobserved heterogeneity at the country level, where, for instance, certain countries tend to attract both foreign investment and immigration, we consider controlling for host-country fixed effects. For instance, certain countries may tend to attract both foreign investment and immigration, due to factors that we have not controlled for. In Section 4.4 below, we will also employ an instrumental variables strategy, with instruments such as the historical migrant share from 40 years ago as well as contemporaneous immigration policies. If migrant shares were time-invariant, the relationship between migrant networks and FDI could not be jointly estimated with host country fixed effects due to perfect collinearity. While migrant shares are not completely time invariant, recall that the variable is only available in 2000, 2005, 2010, and 2013 due to data constraints. We have carried forward the value of migrant share where data is missing, which means there is no time variation within each interval. To address this problem, instead of including fixed effects directly into the estimating equation, we use a proxy, following, for example, Crozet and Hinz (2016). Specifically, using the UN

aggregate bilateral FDI data from 2003 to 2012 for all host and source countries available, we first estimate the following equation:

$$\begin{aligned} \log(FDI_{ijt} + 1) = & b_0 + b_1 \log(GDP_{i,t-1}) + b_2 \log(GDP_{j,t-1}) + \alpha_3 \log(Dist) & (2) \\ & + \alpha_4 Migrant\ share_{i,t-1} + \alpha_5 Migrant\ share_{j,t-1} \\ & + \alpha_X \mathbf{X}_{ij,t-1} + c_i + c_j + c_t + \epsilon_{ijt}. \end{aligned}$$

Therefore, we rely on FDI data not only from China, but all bilateral pairs, and *Migrant share*<sub>*i,t-1*</sub> and *Migrant share*<sub>*j,t-1*</sub> are the Chinese migrant shares of source country *i* and host country *j*, respectively. The estimated coefficients,  $\hat{c}_j$ , provide proxies for host country fixed effects in the regression where only Chinese outward FDI is examined. Results are shown in Table 3 Panel C. Compared to the baseline regression estimates, the coefficients on *Migrant share* are remain positive in all columns, and are quantitatively similar and statistically significant for brownfield FDI. Thus, the gap of the estimated effect for M&A and greenfield actually widens after accounting for host country unobserved heterogeneity.

In Appendix Table 4, we confirm the results are qualitatively similar when host country fixed effects are simply included in the regression, despite the low within-country variation in migrant shares. We use both OLS and the Hausman-Taylor estimator, a commonly used method for estimating the effects of time-invariant regressors.<sup>19</sup> The marginal effects are positive and those on M&A remain larger.

To address endogeneity arising from omitted variable bias, in Table 4, we include in our regression equation measures of growth, financial development, institutional quality, and trade openness. 38 countries are lost as a result.<sup>20</sup> However, as Table 4 shows, the inclusion of these previously omitted variables reveals even stronger differences for greenfield and brownfield FDI. With this specification, the null hypothesis that the *Migrant share* coefficient in columns 1 and 3 are equal is rejected with  $\chi^2$  statistic (*p*-value) of 9.98 ( $< 0.01$ ), and likewise, for columns 2 and 4, the Wald test  $\chi^2$  statistic (*p*-value) is 34.4 ( $< 0.01$ ).

In Table 4, while the coefficients of the control variables are not statistically significant across all of the specifications, some general results emerge, and are consistent with the literature. The estimates suggest that larger countries with lower GDP per capita receives less greenfield investment and fewer M&A with Chinese companies (column 4). This finding is consistent with Chinese multinationals' strategy of going abroad –“encircling the cities from the rural areas”– a pragmatic business strategy that calls for building capacity and the accumulation of wealth in markets with low competition (rural areas) first before moving to developed markets to under-

---

<sup>19</sup>We limit the sample to the years in which the migrant stock data is available to generate larger within-country variation.

<sup>20</sup>Other potential factors include proximity to large markets, and technology or natural resource seeking motives. In unreported results, we confirm that our findings hold qualitatively when we further augment the list of control variables by including measures of remoteness (GDP weighted distance), technology (number of patent applications from the WDI), and natural resource abundance (agricultural raw materials, fuel, and ores and metals exports as a share of merchandise exports from the WDI) (e.g., Huang and Wang, 2013). The sample is substantially reduced to 666 observations and 79 countries in these regressions.

Table 4: Robustness Checks with Additional Regressors

Dependent variable	log (FDI + 1)			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	11.937** (5.608)	32.823 (25.807)	35.087*** (10.658)	181.961*** (42.853)
(log) GDP	0.364*** (0.038)	1.813*** (0.134)	0.128*** (0.029)	0.590*** (0.130)
(log) GDP per capita	-0.266*** (0.054)	-1.287*** (0.255)	-0.081** (0.035)	-0.104 (0.222)
(log) Distance	0.004 (0.090)	-0.519 (0.432)	0.021 (0.076)	-0.075 (0.470)
Language	0.177 (0.220)	1.893* (1.072)	-0.189 (0.411)	-0.560 (1.803)
Legal system	0.126 (0.123)	0.647 (0.601)	0.053 (0.068)	0.001 (0.425)
GDP growth	0.000 (0.002)	0.016 (0.017)	0.001 (0.002)	0.011 (0.014)
Financial development	0.002 (0.001)	0.003 (0.004)	0.003** (0.001)	0.010** (0.005)
Institution quality	0.032 (0.025)	0.111 (0.117)	0.014 (0.017)	-0.001 (0.099)
Trade openness	0.132 (0.080)	0.821** (0.372)	-0.149 (0.101)	-0.854* (0.450)
Year fixed effects	Y	Y	Y	Y
Number of countries	96	96	96	96
Observations	1,021	1,021	1,021	1,021
R-squared	0.575	0.466	0.459	0.331

The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), and likewise for M&A ( $\#BF$  and  $BF$ ). All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.

take competition (cities).<sup>21</sup> Furthermore, markets with higher financial development generally attract more Chinese OFDI. We utilize this full set of control variables for the remainder of our analysis.

Thus far, we have used the Chinese migrant population as a proportion of destination's total population as a measure of the migrant network size. However, it may be that the absolute size of the migrant stock matters. Therefore, in Appendix Table 5, we replace the share of Chinese migrants with the stock Chinese migrant population (in logarithms). This change has no qualitative effect on the positive relationship between migrant networks and outward FDI, and the coefficients also remain statistically significant. A 1 percent increase in the stock of Chinese migrants abroad raises the number (value) of greenfield and brownfield projects by 0.08 and 0.13% (0.17 and 0.33%), respectively.

<sup>21</sup>“Encircling the cities from the rural areas” was initially a military strategy developed by Mao Zedong, the founding Chairman of China. Guided by this strategy, the Communist party established revolutionary bases in rural areas that were largely ignored by the Kuomintang party and gradually accumulated arms forces and wealth to fight with Kuomintang party in cities. This strategy is thought to be crucial for the victory of the Communist party in the domestic war. The strategy also provides important guidelines for business practices, see for example, <http://english.cctv.com/2016/07/11/VIDEaFG3eAExf0417rKdqkGV160711.shtml>

## 4.4 Instrumental variables strategy

### 4.4.1 Historical migrant shares and contemporaneous immigration policy

So far, we have documented a significant and positive relationship between overseas Chinese migrant networks and Chinese OFDI. Moreover, across all specifications, such a linkage is consistently stronger for brownfield investment or cross-border M&A in comparison to greenfield investment. However, one may be concerned of potential endogeneity issues between migration and cross-border investment. As discussed by Javorcik et al. (2012), reverse causality may exist where FDI drives migration instead. For instance, the presence of multinationals from cross-border investment generates greater economic activity, possibly encouraging the inflow of migrants. More directly, overseas migrant networks may be formed by FDI, as employees from the home country are transferred to the host country, and this in turn facilitates the movement of more emigrants to that location. To address this concern, we employ an instrumental variables (IV) approach and use a set of instruments which predict Chinese migrant shares well, thereby meeting the relevance criteria, and also appear to be exogenous with respect to FDI, satisfying the exclusion restriction.

First, following the approach of Javorcik et al. (2012), we use a measure of the historical Chinese migrant network as an instrument. As Javorcik et al. (2012) explain, migration is likely to be correlated over time, as families reunite and established networks in a foreign country lower the cost of immigration. Therefore, this creates a strong pull factor for future migration. We expect large Chinese migrant populations to have the effect of encouraging future migration. Because our primary database for Chinese migrants abroad, the UN Global Bilateral Migration Database, is available from 1990, the longest lag that can be created is 10 years. In this case, there may be concerns that the exclusion restriction for a valid instrument is not satisfied. Hence, we supplement our data with the World Bank Global Bilateral Migration Database, which has data from 1960 to 2000 in 10 year intervals. While the two datasets are not identical, the correlation between them is extremely high for the overlapping years of 1990 and 2000 at 0.977. Therefore, our first instrument is the share of Chinese migrants in 1960 and 1970, essentially a 40-year lag, which is longer than the 30-year lag employed in Javorcik et al. (2012). Because this supplementary data is available every decade, we split our sample between 2003 and 2014 in half and assign the 1960 values for the first half of the sample, and 1970 values for the second half.

In addition to exploiting the variation from historical migrant shares, we also employ contemporaneous instruments that influence migration patterns but not the movement of capital. In particular, we utilize variables which reflect countries' immigration policies purely. Specifically, we first obtain data on whether countries recognize dual citizenship from the MACIMIDE Global Expatriate Dual Citizenship Dataset (Vink et al., 2015). Second, we draw data from a new dataset on the strictness of immigration policies from IMPIC for 33 OECD countries from 1980 to 2010 (Helbling et al., forthcoming). The index employed covers policies from three different fields: family reunification, asylum and refugees, and control of immigration.<sup>22</sup> Thus,

---

<sup>22</sup>The original data contains two additional fields: labor migration and co-ethnics. We exclude these fields

the immigration strictness variable ranges from 0 to 1.

Lastly, since immigration policies may not fully reflect how welcoming a nation is towards immigrants, we use data from the World Values Survey to measure countries' hostility towards foreign residents. The Survey includes a question which asks respondents whether they "would not like to have as neighbors: Immigrants/foreign workers", and is available for three waves of the survey (1999, 2005, and 2010) and a total of 87 countries. Thus, the variable *Immigrant hostility* captures the average hostility towards immigrants for each country, as a fraction between 0 and 1.

Tables 5 reports the two-stage least squares (2SLS) estimation results, with a different set of instruments in each panel. The dependent variable in the first-stage regression is always the Chinese migrant share in the host country, *Migrant share*, and the regressors are the instrument(s), the full set of control variables, and year fixed effects. Note that the first-stage regression as well as estimating sample are identical regardless of the second stage dependent variable. As indicated in column 1 across all panels, the chosen instruments are all significantly correlated with the endogenous variable. For instance, the historical migrant network is a strong positive predictor of the migrant share today, despite it reflecting the share of Chinese migrants roughly 40 years ago. Dual citizenship is also positively correlated with the current share of Chinese migrants. This suggests that dual citizenship is a pull factor on all origin countries, since Chinese citizens cannot accrue its benefits because they must renounce their Chinese nationality if they wish to be a citizen of a foreign country. Not surprisingly, more strict immigration policies are associated with less Chinese immigrants. The estimate in Panel D also shows that the hostility of host countries' residents towards immigrants is negatively correlated with migrant shares, and this effect is statistically significant.

The second-stage IV results in columns 2 to 5 corroborate the OLS estimates and the previous robustness checks. The second-stage results reveal overseas Chinese migrant networks are a strong determinant of Chinese outward FDI. Importantly, the strong positive effects for M&A (i.e., columns 4 and 5) persist, and reinforce the notion that migrants facilitate transactions for the more information-intensive brownfield cross-border investment. Comparing the estimates on *Migrant share* from Table 1 with Table 5 reveal that the OLS coefficients are actually quite close to the IV estimates.

The diagnostic tests also indicate that, in general, the instruments chosen are valid. The  $p$ -values associated with the Kleibergen-Paap rank Lagrange multiplier statistic from the underidentification test and the Hansen J statistic from a test of overidentifying restrictions are reported. For the former,  $p$ -values are all very small, rejecting the null hypothesis that the excluded instruments are not relevant. For the latter, the  $p$ -values are generally greater than 0.1, suggesting that the exclusion restriction is satisfied and the instruments are uncorrelated with the error term. Only in Tables 5 Panel D column 2, the  $p$ -value is slightly less than 0.1 at 0.07.

---

because the former relates to existing labor market conditions, such as the quality of migrants accepted, while the latter is already captured in part by historical migrant shares.

Table 5: IV Estimation with Two-stage Least Squares- 40-year Lagged Migrant Shares

Panel A					
Dependent variable:	Stage I	Stage II			
	Migrant share (1)	# GF (2)	GF (3)	# BF (4)	BF (5)
Migrant share (40-year lag)		13.572*** (4.355)	47.908** (20.698)	22.875*** (3.822)	123.934*** (17.023)
Migrant share	0.381*** (0.023)				
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Observations	1,007	1,007	1,007	1,007	1,007
R-squared	0.819				
Panel B					
Dependent variable:	Stage I	Stage II			
	Migrant share (1)	# GF (2)	GF (3)	# BF (4)	BF (5)
Migrant share (40-year lag)	0.382*** (0.023)				
Dual citizenship	0.001*** (0.000)				
Migrant share		13.292*** (4.342)	50.534** (20.763)	22.714*** (3.823)	123.457*** (17.029)
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Overidentification test ( $p$ )		0.34	0.12	0.47	0.76
Observations	1,007	1,007	1,007	1,007	1,007
R-squared	0.820				
Panel C					
Dependent variable:	Stage I	Stage II			
	Migrant share (1)	# GF (2)	GF (3)	# BF (4)	BF (5)
Migrant share (40-year lag)	5.900*** (0.382)				
Dual citizenship	-0.0001 (0.0001)				
Immigrant strictness	-0.001 (0.001)				
Migrant share		2.332 (21.807)	151.573 (112.018)	118.521*** (18.993)	666.083*** (89.383)
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Overidentification test ( $p$ )		0.53	0.43	0.28	0.35
Observations	266	266	266	266	266
R-squared	0.883				
Panel D					
Dependent variable:	Stage I	Stage II			
	Migrant share (1)	# GF (2)	GF (3)	# BF (4)	BF (5)
Migrant share (40-year lag)	0.421*** (0.042)				
Dual citizenship	0.002*** (0.000)				
Immigrant hostility	-0.003** (0.001)				
Migrant share		4.752 (7.123)	73.707*** (27.278)	22.224*** (7.791)	115.310*** (29.924)
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Overidentification test ( $p$ )		0.07	0.59	0.91	0.76
Observations	378	378	378	378	378
R-squared	0.844				

*Notes:* In all panels, the dependent variable of the first stage regression is *Migrant share*. The second stage dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), and likewise for M&A ( $\#BF$  and  $BF$ ). All regressions include the full set of control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively. The  $p$ -values associated with the Kleibergen-Paap rank Lagrange Multiplier (Hansen J) statistic for the underidentification (overidentification) test are reported.

#### 4.4.2 Predicted migrant shares

China's open-door policy beginning in 1978 marked a dramatic change in economic policy, leading to its eventual membership in the World Trade Organization in 2001. As Figure 1 makes clear, China's OFDI only rises significantly at the turn of the 21st century. However, for various historical reasons, the overseas Chinese community had always been sizable, and so the size of the network 40 years ago would be difficult to attribute to cross-border movements in capital. Nonetheless, to further alleviate concerns regarding the validity of this particular instrument, we present an alternative to lagged migrant shares. Specifically, we take advantage of the bilateral structure of the data and construct an instrumental variable following the logic of Burchardi et al. (2016).<sup>23</sup>

The method utilizes information on global migration outflows and inflows which capture the push and pull factors of migration, respectively. First, we predict the stock of migrants in a particular destination  $j$  if country  $j$  received an inflow of Chinese immigrants (in percentage terms) equal to the rest of the world. Second, we adjust this inflow based on the proportion of non-Chinese immigration relative to the rest of the world, such that countries which receive more (non-Chinese) immigrants overall will also experience a larger influx of Chinese migrants. Denote  $Migrant_{j,t}^{CHN}$  as the stock of Chinese migrants in country  $j$  time  $t$ . Then, the first step computes the push factor as the percentage change in the flow of Chinese migrants to all destinations other than  $j$ :

$$Push_{j,t} = \frac{Migrant_{-j,t}^{CHN}}{Migrant_{-j,t-1}^{CHN}} - 1 \equiv \frac{Migrant_{WLD,t}^{CHN} - Migrant_{j,t}^{CHN}}{Migrant_{WLD,t-1}^{CHN} - Migrant_{j,t-1}^{CHN}} - 1$$

This global outflow of Chinese migrants captures the push factor that does not depend on factors in destination  $j$ . In the second step, the push factor is adjusted by the inflow of immigrants relative to the rest of the world by the factor:

$$Pull_{j,t} = \left( \frac{Migrant_{j,t}^{-CHN}}{Migrant_{j,t-1}^{-CHN}} - 1 \right) \div \left( \frac{Migrant_{WLD,t}^{-CHN}}{Migrant_{WLD,t-1}^{-CHN}} - 1 \right)$$

For example, if the pull factor is larger than 1, then the increase in foreign population in destination  $j$  is larger than the global rise in migrant population. Combining these two steps, the predicted migrant stock in destination  $j$  time  $t$  is then

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

In this step, the foreign population inflow into the US originating outside of China (relative to the world) captures the pull factor for which China is irrelevant. If the push factor is negative,

<sup>23</sup>We found that applying the method of Burchardi et al. (2016) directly here yields extreme predictions of migrant stocks. This is in part due to the decreasing stocks observed in countries, driving some migrant stocks, which were already small to begin with, to be negative.

it is divided by the pull factor because the destination country should be losing fewer Chinese migrants if it attracts more immigrants in general. Since 1960 is the earliest year available, the actual migrant stock is used as opposed to the predicted, i.e.,  $\widehat{Migrant}_{j,1960}^{CHN} = Migrant_{j,t}^{CHN}$ . We then repeat this calculation for all destinations and for every interval in which migration data is available (i.e., 1970, 1980, 1990, 1995, 2000, 2005, and 2010). The predicted values for the years 2000, 2005, and 2010, divided by the destination country's population in the respective years, are used as instruments for *Migrant share*.

Taking the US from 1960-70 as an example, the Chinese-born population of US in 1960 and 1970 are 105,384 and 220,531, respectively. The global stock of Chinese migrants in 1960 and 1970 are 4,803,292 and 4,448,338, respectively. Hence, the push and pull factors are equal to

$$Push_{US,1970} = \frac{4,448,338 - 105,384}{4,803,292 - 220,531} - 1 = -0.100$$

$$Pull_{US,1970} = \left( \frac{11,973,797 - 220,531}{10,825,585 - 105,384} - 1 \right) \div \left( \frac{105,509,001 - 4,448,338}{92,825,210 - 4,803,292} - 1 \right) = 0.651$$

If the growth rates of immigrants in the US is larger than the world, then we predict the growth rate of Chinese immigrants should be as well. After this proportional adjustment, the predicted migrant stock for 1970 is

$$\widehat{Migrant}_{US,1970}^{CHN} = 105,384 \times \left( 1 - \frac{0.100}{0.651} \right) = 89,174.$$

Table 6 uses these predicted migrant shares as well as the previous contemporaneous variables as IVs. The predicted migrant share is strongly correlated with the actual migrant share. For the most part, the coefficients in Tables 5 and 6 are quantitatively similar, and the second-stage results reinforces our earlier findings with OLS. Again, the diagnostic tests perform well in all the panels.

## 4.5 Heterogeneous effects

Having established the positive link between migrant networks and FDI, we now explore the different channels through which overseas migrants might facilitate cross-border investment. In doing so, we seek to understand why the relationship is stronger for M&A as opposed to greenfield investment. Thus, we present empirical evidence showing a more pronounced migrant network effect for FDI facing greater informal barriers, in particular, when information intensity is higher, as well as formal barriers with regards to countries' regulations. With the detailed deal-level FDI data by mode of entry, we exploit the cross-sectional variation of investment at the industry and country level, and compare FDI from privately-owned enterprises (POE) versus state-owned enterprises (SOE).



Table 6: IV Estimation with Two-stage Least Squares- Predicted Migrant Shares

Panel A					
Dependent variable:	Stage I	Stage II			
	Migrant share	# GF	GF	# BF	BF
	(1)	(2)	(3)	(4)	(5)
Predicted migrant share		14.696*** (3.505)	40.490** (18.869)	21.930*** (3.876)	135.867*** (22.255)
Migrant share	0.136*** (0.017)				
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Observations	982	982	982	982	982
R-squared	0.787				
Panel B					
Dependent variable:	Stage I	Stage II			
	Migrant share	# GF	GF	# BF	BF
	(1)	(2)	(3)	(4)	(5)
Predicted migrant share	0.137*** (0.017)				
Dual citizenship	0.001*** (0.000)				
Migrant share		14.346*** (3.532)	43.630** (19.100)	21.739*** (3.881)	134.940*** (22.300)
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Overidentification test ( $p$ )		0.36	0.16	0.51	0.64
Observations	982	982	982	982	982
R-squared	0.789				
Panel C					
Dependent variable:	Stage I	Stage II			
	Migrant share	# GF	GF	# BF	BF
	(1)	(2)	(3)	(4)	(5)
Predicted migrant share	0.274*** (0.045)				
Dual citizenship	0.001*** (0.0004)				
Immigration strictness	-0.005*** (0.002)				
Migrant share		-12.570 (36.642)	212.230 (165.623)	72.598** (29.236)	443.436** (198.251)
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Overidentification test ( $p$ )		0.49	0.40	0.23	0.20
Observations	266	266	266	266	266
R-squared	0.420				
Panel D					
Dependent variable:	Stage I	Stage II			
	Migrant share	# GF	GF	# BF	BF
	(1)	(2)	(3)	(4)	(5)
Predicted migrant share	0.102*** (0.019)				
Dual citizenship	0.002*** (0.000)				
Immigrant hostility	-0.009*** (0.002)				
Migrant share		4.049 (8.020)	80.245** (38.732)	25.091*** (8.172)	121.298*** (43.195)
Underidentification test ( $p$ )		< 0.01	< 0.01	< 0.01	< 0.01
Overidentification test ( $p$ )		0.07	0.60	0.90	0.73
Observations	378	378	378	378	378
R-squared	0.780				

*Notes:* In all panels, the dependent variable of the first stage regression is *Migrant share*. The second stage dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), and likewise for M&A ( $\#BF$  and  $BF$ ). All regressions include the full set of control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively. The  $p$ -values associated with the Kleibergen-Paap rank Lagrange Multiplier (Hansen J) statistic for the underidentification (overidentification) test are reported.

#### 4.5.1 Heterogeneity across industries

First, we examine heterogeneity across different industries. While roughly half of global foreign investment, whether greenfield or brownfield, is made in the tertiary or service industries, the problem of asymmetric information is perhaps especially severe for such cross-border transactions.<sup>24</sup> With their know-how and comparative advantage in combining resources from both sides of the border more efficiently, overseas Chinese are expected to add more value to industries with greater knowledge-intensity, such as information technology, engineering, and research and development (Coff, 1999).<sup>25</sup> Migrants are thus expected to mitigate informal problem associated with information as well as more formal barriers.

To aggregate the total number and value of greenfield and brownfield projects to the industry level, we rely on Orbis to provide the industry of the parent company for greenfield FDI, and SDC Platinum itself for the acquirer's industry. We classify industries as either primary (SIC two-digit code 1-14), secondary (15-39), or tertiary (40-99). Primary industries are related to agriculture or natural resources, secondary industries are generally manufacturing, and tertiary industries include services for the most part.<sup>26</sup> The number or value of FDI projects is aggregated up to this broader industry definition, i.e.,  $s = \text{primary, secondary, or tertiary}$ , in the following specification:

$$\begin{aligned} \log(FDI_{sjt} + 1) = & g_0 + g_1 \log(GDP_{j,t-1}) + g_2 \log(Dist) + g_3 Migrant\ share_{j,t-1} \\ & + g_4 Migrant\ share_{j,t-1} \times Tertiary_s + g_5 Tertiary_s + \gamma \mathbf{X}_{j,t-1} + v_{sjt}, \end{aligned} \quad (3)$$

where  $Tertiary_s$  is an indicator variable equal to one if  $s = \text{tertiary}$ , and zero otherwise.<sup>27</sup> In Table 7 Panel A, we see that the largest effects are indeed observed for the knowledge-intensive tertiary industries. Note that even within an industry, the positive relationship is more pronounced for M&A. Thus, consistent with our hypothesis, the evidence suggests that the positive relationship between overseas Chinese networks and Chinese OFDI is mostly driven by the investment in the tertiary or service industries.<sup>28</sup>

While knowledge-intensity may vary across sectors, a company that invests outside the industry it operates in may also find the environment to be less familiar and encounter information barriers. Thus, the migrant network effect is expected to be larger for FDI in industries that are different from that of the investing company. This hypothesis is tested in Table 7 Panel B. For greenfield investment, while the SIC code for the parent company can be obtained, the

<sup>24</sup>See Annex Tables 14 and 20 from (UNCTAD, 2016)

<sup>25</sup>A report by the World Bank (2010) also found fewer restrictions on foreign ownership in primary and manufacturing sectors, but stricter limits in services, which reinforces the migrant effect.

<sup>26</sup>Specifically, primary industries are Agriculture, Forestry, Fishing, Mining; secondary industries are Construction and Manufacturing; and tertiary industries are Transportation, Communications, Electric, Gas, Sanitary Services, Wholesale Trade, Retail Trade, Finance, Insurance, Real Estate, Services, and Public Administration.

<sup>27</sup>In unreported results, we also included an additional indicator variable for secondary industries and its interaction with *Migrant share*. The results are qualitatively and quantitatively similar.

<sup>28</sup>In unreported results, we confirm that when the baseline specification is estimated for each group of industries, the migrant network effect is not statistically different from zero for primary industries, while the effect is positive for secondary industries, and largest for tertiary industries.

Table 7: Heterogeneity Across Industries

	Panel A: Tertiary versus Primary and Secondary			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	-5.639 (4.804)	-39.825 (24.906)	15.373** (6.212)	102.938*** (36.871)
Migrant share $\times$ Tertiary	17.964*** (6.327)	111.715** (50.539)	14.755*** (2.496)	70.678*** (7.155)
Tertiary	-0.261*** (0.055)	-2.350*** (0.411)	-0.042 (0.039)	-0.433* (0.223)
Observations	1,513	1,513	1,513	1,513
R-squared	0.408	0.376	0.344	0.240

	Panel B: Cross-Industry FDI			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	7.013 (4.597)	21.914 (25.414)	22.034*** (7.443)	129.555*** (41.234)
Migrant share $\times$ Cross-industry	2.066** (0.881)	13.244** (5.102)	9.883*** (1.825)	68.656*** (5.601)
Cross-industry	0.061** (0.023)	0.095 (0.165)	0.020 (0.015)	0.014 (0.120)
Observations	1,770	1,770	1,600	1,600
R-squared	0.458	0.373	0.373	0.270

*Notes:* The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), and likewise for M&A ( $\#BF$  and  $BF$ ). All regressions include the full set of control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.

industry to which the FDI project itself is not coded. We assign the project a SIC 2-digit code based on information provided from fDi Intelligence regarding the industry sector, sub-sector, and industry activity. If this code is different from that of the parent company, we refer to this as cross-industry FDI. Meanwhile, brownfield FDI is classified as cross-industry if the industry code of the acquirer and target companies are different. Similar to the previous specifications, aggregate FDI at the country level is computed separately for investment in the same industry and across different industries (i.e.,  $s$  = same industry or cross-industry). Subsequently, an indicator for cross-industry FDI is interacted with *Migrant share*. As Table 7 Panel B shows, consistent with idea that migrants promote FDI through an information channel, the marginal effect of the size of migrant networks on FDI is more pronounced for cross-industry FDI.

#### 4.5.2 Heterogeneity across host countries

Next, we consider how the effect of migrant networks varies with host-country characteristics. In particular, we provide further evidence that emigrant networks matter more when the barriers to international transactions, whether formal or informal, are larger. The results of this analysis are presented in Table 8. First, in Panel A, in addition to the full set of control variables, we include an interaction term between *Migrant share* and the common language indicator variable. The estimates show that for both greenfield and brownfield OFDI, the effect

Table 8: Heterogeneity across Host Countries- Barriers to Entry

	Panel A			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	44.228*	205.689**	116.789***	523.288***
	(26.263)	(89.564)	(29.248)	(90.869)
Migrant share $\times$ Language	-38.503	-207.433**	-96.864***	-405.228***
	(26.149)	(90.948)	(28.972)	(91.049)
Language	0.362***	2.833***	0.297***	1.451***
	(0.111)	(0.547)	(0.077)	(0.445)
Observations	1,019	1,019	1,019	1,019
R-squared	0.581	0.474	0.538	0.370
	Panel B			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	5.503	4.623	24.446***	142.969***
	(3.477)	(12.188)	(3.759)	(21.467)
Migrant share $\times$ Capital controls	58.967***	240.422***	128.723***	466.271***
	(15.731)	(86.096)	(16.632)	(80.399)
Capital controls	-0.056	-0.378	-0.133***	-0.573**
	(0.076)	(0.370)	(0.041)	(0.243)
Observations	970	970	970	970
R-squared	0.590	0.478	0.547	0.364
	Panel C			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	7.046*	15.308	28.426***	158.351***
	(4.057)	(17.550)	(6.510)	(29.348)
Migrant share $\times$ Labor regulations	34.979**	147.715**	85.114***	317.872***
	(13.964)	(70.344)	(18.498)	(82.100)
Labor regulations	-0.081	-0.263	-0.160**	-0.456
	(0.079)	(0.403)	(0.066)	(0.293)
Observations	907	907	907	907
R-squared	0.585	0.464	0.502	0.344
	Panel D			
	# GF (1)	GF (2)	# BF (3)	BF (4)
Migrant share	8.732*	17.644	29.761***	163.608***
	(4.568)	(17.805)	(7.199)	(32.506)
Migrant share $\times$ Administrative requirements	16.531	134.696*	78.037***	269.166**
	(12.249)	(78.602)	(23.357)	(110.062)
Administrative requirements	0.035	0.255	-0.140*	-0.180
	(0.070)	(0.383)	(0.071)	(0.325)
Observations	907	907	907	907
R-squared	0.583	0.465	0.492	0.340

*Notes:* The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), and likewise for M&A ( $\#BF$  and  $BF$ ). All regressions include the full set of control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.

of migrants is larger for countries that only have a small percentage of its population speaking Chinese, i.e., when *Language* is equal to 0. This contrast is stronger for M&A. Since sharing a common language implies a lower language barrier, smaller cultural gap, and greater trust between the countries, migrants' role in facilitating international investment transactions is expected to shrink. Indeed, the empirical evidence validates this hypothesis.

To provide further insight on the role of migrants in overcoming formal regulation barriers, we employ a similar method to examine the heterogeneous effects across countries with different regulations and laws. From the Economic Freedom of the World database of (Gwartney et al., 2015), we obtain measures of capital controls, labor regulations involved with hiring and firing employees, and business regulations from administrative requirements issued by the government that companies have to comply with.<sup>29</sup> To contrast countries with tight versus loose regulations, we convert each of these variables to indicator variables that take the value of 1 if the measure of the country's barrier is stronger than the median value across the sample, and zero otherwise. Again, the share of Chinese migrants is interacted with these dummy variables. When regulations on the international flow of capital, labor market, business practices are higher, migrants have a stronger role in increasing the OFDI from their origin country, especially for M&A.<sup>30</sup>

### 4.5.3 Heterogeneity across firms

First, we test the hypothesis that when information about the investing or invested company is more scarce, the marginal effect of migrant networks on FDI increases. Analyzing heterogeneous effects at the firm level can reveal the potential benefits accrued to individual companies from utilizing the international connections. Companies can be classified by whether or not they are listed on a stock exchange. While measures of a company's transparency are difficult to obtain, the general public typically has more information about a company that is listed or publicly traded. For instance, besides potentially greater name recognition, the listed company must disclose certain financial details and file earnings reports. As argued by Erel et al. (2012), information asymmetry is likely to be a larger problem for private targets as opposed to public targets for cross-border M&As. Thus, we expect migrants to have a larger role in facilitating FDI for privately held (i.e., unlisted) as opposed to publicly traded companies .

To identify whether the parent company of greenfield FDI is a private or publicly traded firm, we match our deal-level data to Datastream. For M&A, we consider the status of both the acquirer from China and the target in the host country. Following Erel et al. (2012), an

---

<sup>29</sup>Gwartney et al. (2015) compiles data from various sources. The International Monetary Fund, Annual Report on Exchange Arrangements and Exchange Restrictions reports on up to 13 types of international capital controls. The zero-to-10 rating is the percentage of capital controls not levied as a share of the total number of capital controls listed, multiplied by 10. Hiring and firing regulations and administrative requirements are both based on questions from the World Economic Forum Global Competitiveness Report. The questions, respectively, are "The hiring and firing of workers is impeded by regulations (= 1) or flexibly determined by employers (= 7)." and "Complying with administrative requirements (permits, regulations, reporting) issued by the government in your country is (1 = burdensome, 7 = not burdensome)."

<sup>30</sup>Note that the magnitude of the coefficients on *Migrant share* can become quite large because of the split in sample: the maximum value of *Migrant share* for countries that have strict capital controls is 1.9%, while for those that do not, it is 7.2%.

acquirer or target firm is listed if its public status is “Public” or non-missing. We separately aggregate to the country level the values of our dependent variables (i.e.,  $\#GF$ ,  $GF$ ,  $\#BF$ , and  $BF$ ) for the different statuses,  $s = \text{private or public}$ . Hence, the following equation is estimated at the country-status level:

$$\begin{aligned} \log(FDI_{s jt} + 1) = & \gamma_0 + \gamma_1 \log(GDP_{j,t-1}) + \gamma_2 \log(Dist) + \gamma_3 Migrant\ share_{j,t-1} \\ & + \gamma_4 Migrant\ share_{j,t-1} \times Private_s + \gamma_5 Private_s + \gamma_X \mathbf{X}_{j,t-1} + \nu_{s jt}, \end{aligned} \quad (4)$$

where  $Private_s$  is an indicator variable equal to 1 if the FDI is associated with private companies and zero otherwise. If the coefficient  $\gamma_4$  is positive, this would corroborate the idea that overseas migrants can help resolve the barrier of information asymmetry. The estimation results are presented in Table 9 Panel A, where in columns 1 and 2, greenfield FDI originating from private parent companies is considered; in columns 3 and 4 (5 and 6), the link between migrant networks and brownfield FDI from private acquirers (to private targets) is examined. As before, the effect on brownfield investment remains larger than on greenfield FDI (whether by count or value), regardless of whether the investment is originating from a firm that is private or publicly traded. Moreover, as hypothesized,  $\gamma_4$  is positive and statistically significant across all columns, suggesting migrants are more important in facilitating cross-border investment when less information is available about either one of the involved parties. Lastly, we contrast the relationship between migrant networks and FDI from privately-owned enterprises (POE) versus state-owned enterprises (SOE). Indeed, the business practices of firms can vary widely. In particular, according to a survey on Chinese outward direct investment by the China Council for the Promotion of 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 shows that 93.5% POEs highly value opinions from their overseas employees while this number is 80.7% for SOEs (CCPIT, 2015, Figure 6.8). This evidence suggests that POEs are more open to utilizing overseas Chinese networks than SOEs. However, because SOEs are affiliated with foreign governments, their investment deals are more heavily scrutinized and may face higher entry barriers and greater restrictions from the host-country.<sup>31</sup> Thus, whether migrants play a larger role for state-owned or privately-owned enterprises is theoretically ambiguous.

In Table 9 Panel B, we run a similar regression to Equation (4), but instead of status, FDI is aggregated up to the country level depending on different ownership types, POE or SOE. Hence,  $Private_s$  is replaced by an indicator for POEs. Empirically, migrant shares have a stronger relationship to investment from private companies as opposed to those with government support. The coefficient on the interaction term is positive and statistically significant, corroborating the idea that companies without government support can utilize the networks abroad to help overcome barriers associated with international transactions, whether they are informal, related

<sup>31</sup>For example, in 2012, former Canadian Prime Minister Stephen Harper approved the sale of Canada’s oil company Nexen to China’s state-owned energy giant China National Offshore Oil Corporation (CNOOC), but restricted foreign SOEs to minority stakes in the future except in “exceptional circumstances”. (See “Canada OK’s foreign energy takeovers, but slams door on any more”, *Reuters*, Dec. 8, 2012.)

Table 9: Heterogeneity Across Firms

Private company	Panel A: Private versus Public Firms					
	Parent		Acquirer		Target	
	# GF (1)	GF (2)	# BF (3)	BF (4)	# BF (5)	BF (6)
Migrant share	5.583 (4.153)	24.311 (24.778)	20.192*** (5.861)	136.454*** (37.809)	22.494*** (8.133)	159.693*** (42.367)
Migrant share $\times$ Private	5.382** (2.167)	20.750*** (6.776)	11.381** (5.607)	42.899** (19.210)	8.036*** (2.195)	15.346*** (5.422)
Private	0.354*** (0.041)	2.194*** (0.191)	0.150*** (0.025)	0.802*** (0.159)	0.197*** (0.031)	0.795*** (0.142)
Observations	1,770	1,770	1,600	1,600	1,600	1,600
R-squared	0.468	0.401	0.382	0.268	0.377	0.281

	Panel B: State-owned versus Privately-owned Enterprises			
	# GF (1)	GF (2)	# BF (3)	BF (4)
	Migrant share	6.214 (4.005)	12.696** (6.030)	40.459* (23.155)
Migrant share $\times$ POE	8.076*** (2.189)	20.637*** (5.873)	16.483* (8.574)	80.231*** (18.189)
POE	0.222*** (0.040)	0.098*** (0.028)	0.829*** (0.194)	0.398*** (0.139)
Observations	2,038	2,038	2,038	2,038
R-squared	0.459	0.364	0.389	0.249

*Notes:* The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), and likewise for M&A ( $\#BF$  and  $BF$ ). All regressions include the full set of control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.

to information asymmetry, or formal regulations.

## 5 Conclusion

Using data on Chinese migrant stocks across many countries and Chinese outward FDI, this paper provides empirical evidence that migration and cross-border investment are closely linked. We find migrant networks to be an important determinant of FDI at the aggregate level, as well as the disaggregate level by mode of entry, i.e., greenfield and brownfield FDI. Moreover, effects are observed at both the extensive and intensive margins, as captured by the number and value of investment projects, respectively. Furthermore, consistent with the idea that M&A face greater information asymmetry and other informal and formal barriers, the results indicate overseas Chinese networks are more closely associated with brownfield FDI than greenfield FDI. We also provide evidence that migrant networks facilitate cross-border investment by alleviating information frictions and cultural barriers. The relationship between overseas Chinese networks and OFDI is more pronounced for private as opposed to publicly traded companies, privately-owned as opposed to state-owned enterprises, for the knowledge-intensive tertiary industries compared to secondary and primary industries, for cross-industry FDI, and in host-countries with smaller Chinese-speaking populations and tighter regulations.

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 the greater labor mobility, according to the World Bank World Development Indicators, the global stock of migrants has surged from around 150 to roughly 250 million between 1990 and 2015. Moreover, this has been accompanied by tremendous growth in the cross-border flow of capital: the outward stock of FDI (as a percentage of GDP) for the world has tripled since 1990 (from 10.1% to 34.0%).<sup>32</sup> For countries like India and South Africa, FDI outflows are 2.5 and 6 times larger, respectively, since 2005. Globalization is a trend that is likely to persist in the near future. As barriers in the international movement of labor and capital fall and the global economy becomes even more integrated, understanding their linkages is key in the future research agenda.

---

<sup>32</sup>See Annex Table 8 from (UNCTAD, 2016).



## References

- Beck, T., A. Demirgüç-Kunt, and R. Levine (2000). A new database on the structure and development of the financial sector. *World Bank Economic Review* 14(3), 597–605.
- 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.
- Bevan, A. A. and S. Estrin (2004). The determinants of foreign direct investment into European transition economies. *Journal of Comparative Economics* 32, 775–787.
- 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.
- Blonigen, B. A., R. B. Davies, G. R. Waddell, and H. T. Naughton (2007). FDI in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review* 51, 1303–1325.
- 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.
- Buckley, P. J., L. J. Clegg, A. R. Cross, X. Liu, H. Voss, and P. Zheng (2011). The determinants of Chinese outward foreign direct investment. *Journal of International Business Studies* 38, 499–518.
- Burchardi, K. B., T. Chaney, and T. A. Hassan (2016). Migrants, ancestors, and investment. *NBER Working Paper No. 21847*.
- Burchardi, K. B. and T. A. Hassan (2013). The economic impact of social ties: Evidence from German reunification. *Quarterly Journal of Economics*.
- Chang, P.-l. (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 (2016). Outward FDI and domestic input distortions: Evidence from Chinese firms. *University of Hong Kong - mimeo*.

- 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.
- 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*.
- Coff, R. W. (1999). How buyers cope with uncertainty when acquiring firms in knowledge-intensive industries: Caveat emptor. *Organization Science* 10, 144–161.
- Cohen, L., U. G. Gurun, and C. Malloy (2016). Resident networks and corporate connections: Evidence from World War II internment camps. *Journal of Finance*.
- 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.
- Crozet, M. and J. Hinz (2016). Friendly fire: The trade impact of the russia sanctions and counter-sanctions. *Kiel Institute for the World Economy - mimeo*.
- Davies, R. B., R. Desbordes, and A. Ray (2015). Greenfield versus merger & acquisition fdi: Same wine, different bottles? *UCD Centre for Economic Research Working Paper Series WP15/03*.
- Deng, P. (2004). Outward investment by Chinese MNCs: Motivations and implications. *Business Horizons* 47(3), 8–16.
- Desbordes, R. and S.-J. Wei (2014). The effects of financial development on foreign direct investment. *World Bank Policy Research Working Paper No. WPS 7065*.
- 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.
- Gao, T. (2003). Ethnic Chinese networks and international investment: Evidence from inward FDI in China. *Journal of Asian Economics* 14(4), 611–629.
- 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.
- 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](http://www.freetheworld.com/datasets_efw.html).

- Head, K. and J. Ries (1998). Immigration and trade creation: Econometric evidence from Canada. *Canadian Journal of Economics* 31, 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.
- Helbling, M., L. Bjerre, F. Römer, and M. Zobel (forthcoming). Measuring immigration policies: The impic database. *European Political Science*.
- Hennart, J.-F. and Y.-R. Park (1993). Greenfield vs. acquisition: The strategy of Japanese investors in the United States. *Management Science* 39, 1054–1070.
- 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 (2011). Chinese outward direct investment: Is there a China model? *China & World Economy* 19(4), 1–21.
- 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 (2012). Migrant networks and foreign direct investment. *Journal of Development Economics* 94, 231–241.
- Kolstad, I. and A. Wiig (2012). What determines Chinese outward FDI? *Journal of World Business* 47, 26–34.
- Kugler, M. and H. Rapoport (2007). International labor and capital flows: Complements or substitutes? *Economic Letters* 94(2), 155–162.
- La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny (1999). The quality of government. *Journal of Law, Economics and Organization* 15(1), 222–279.
- Lee, H.-H. (2016). Policy factors influencing FDI inflows to developing countries. *Asian Development Bank Asian Economic Integration Report 2016*.
- Nocke, V. and S. Yeaple (2008). An assignment theory of foreign direct investment. *Review of Economic Studies* 75, 529–557.
- Parsons, C. and P.-L. Vézina (forthcoming). Migrant networks and trade: The vietnamese boat people as a natural experiment. *Economic Journal*.
- Portes, R. and H. Rey (2005). The determinants of cross-border equity flows. *Journal of International Economics* 65, 269–296.
- Qiu, L. D. and S. Wang (2011). FDI policy, greenfield investment and cross-border mergers. *Review of International Economics* 19(5), 836–851.

- Rauch, J. E. (2001). Business and social networks in international trade. *Journal of Economic Literature* 39, 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.
- Santos Silva, J. M. C. and S. Tenreyro (2006). The log of gravity. *Review of Economics and Statistics* 88(4), 641–658.
- Stepanok, I. (2015). Cross-border mergers and greenfield foreign direct investment. *Review of International Economics* 23(1), 111–136.
- Tian, W. and M. Yu (2015). Distribution outward FDI and productivity heterogeneity: Evidence from Chinese firms. *Peking University - mimeo*.
- 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 (2016). *World investment report*. New York.
- Vink, M. P., G.-R. De Groot, and N. C. Luk (2015). MACIMIDE Global Expatriate Dual Citizenship Dataset. <http://dx.doi.org/10.7910/DVN/TTMZ08>.
- Wang, B., Y. Tan, M. Yu, and Y. Huang (2016). Outward direct investment, firm productivity and credit constraints: Evidence from Chinese firms. *Pacific Economic Review* 21, 72–83.
- World Bank (2010). Investing across borders 2010. <http://iab.worldbank.org/~media/FDKM/IAB/Documents/IAB-report.pdf>. Accessed: 08/10/2016.

# Appendix

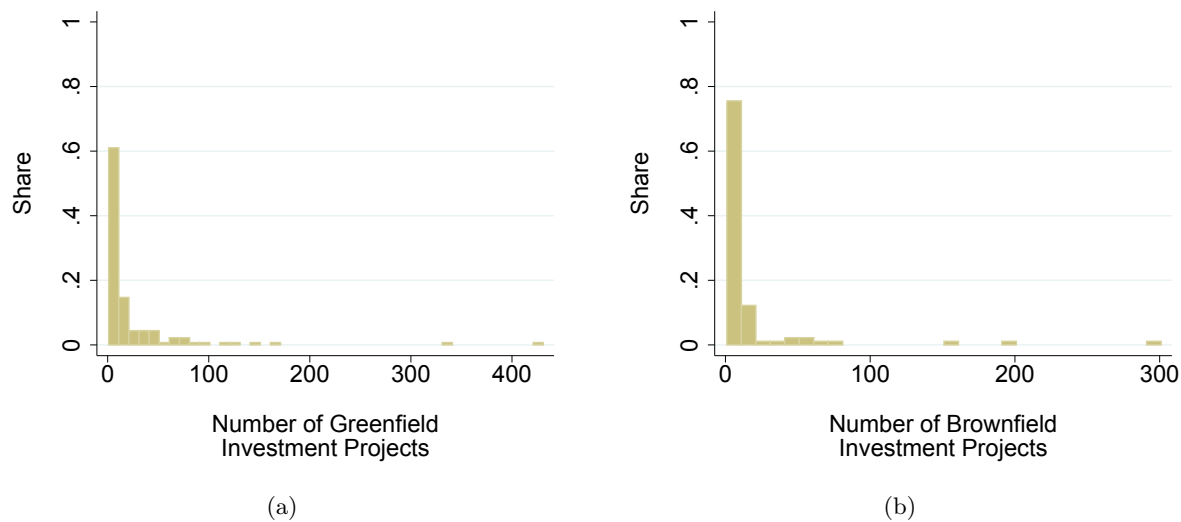
Appendix Table 1: Number of Investment Projects by Country between 2003 and 2014

Both Greenfield (GF) and Brownfield (BF)								
Country	#GF	#BF	Country	#GF	#BF	Country	#GF	#BF
Argentina	14	5	Hungary	29	8	Papua New Guinea	3	1
Australia	77	155	India	127	6	Peru	14	5
Austria	1	6	Indonesia	65	13	Philippines	36	1
Azerbaijan	9	3	Iraq	1	2	Poland	33	1
Belarus	11	1	Ireland	11	2	Portugal	5	3
Belgium	27	4	Israel	6	7	South Korea	36	19
Bolivia	1	2	Italy	37	26	Romania	35	1
Brazil	94	15	Jamaica	4	1	Rusia	111	11
Cambodia	14	1	Japan	63	56	Singapore	78	57
Canada	56	73	Jordan	6	1	Slovakia	4	1
Cayman Islands	1	5	Kazakhstan	14	9	South Africa	48	1
Chile	9	2	Kyrgyzstan	5	2	Spain	41	15
Colombia	9	4	Liberia	1	1	Sri Lanka	3	2
Cyprus	2	1	Lithuania	7	1	Sweden	22	8
Czech Republic	13	3	Luxembourg	7	3	Switzerland	13	6
Dem. Rep. of Congo	5	1	Macau, China	7	3	Syrian Arab Republic	3	1
Denmark	24	7	Malaysia	48	12	Taiwan	73	16
Ecuador	6	2	Mexico	33	4	Thailand	48	11
Egypt	17	3	Mongolia	7	11	Tunisia	2	1
Estonia	1	1	Mozambique	3	1	Turkey	22	4
Finland	3	1	Namibia	1	1	Ukraine	7	7
France	82	35	Netherlands	46	2	United Arab Emirates	43	1
Gabon	2	4	New Zealand	8	11	United Kingdom	161	44
Georgia	2	1	Nigeria	11	2	United States	333	191
Germany	43	61	North Korea	6	1	Viet Nam	64	7
Ghana	9	1	Norway	3	4	Zambia	18	1
Hong Kong, China	149	301	Pakistan	28	4	Zimbabwe	7	2
Only Greenfield (GF)								
Country	#GF	Country	#GF	Country	#GF	Country	#GF	
Afghanistan	3	Greece	13	Paraguay	2			
Algeria	12	Guyana	5	Qatar	5			
Angola	1	Haiti	1	Rwanda	4			
Antigua and Barbuda	1	Honduras	2	Saudi Arabia	12			
Armenia	1	Iran	12	Senegal	3			
Bahrain	3	Kenya	11	Serbia	2			
Bangladesh	6	Kuwait	4	Slovenia	1			
Bosnia and Herzegovina	2	Lao	1	Sudan	4			
Botswana	1	Latvia	1	Tajikistan	5			
Brunei Darussalam	3	Madagascar	2	Macedonia	1			
Bulgaria	16	Micronesia	1	Turkmenistan	5			
Ivory Coast	1	Moldova	1	Uganda	5			
Cameroon	3	Morocco	3	Tanzania	4			
Chad	2	Myanmar	9	Uruguay	3			
Costa Rica	1	Nepal	2	Uzbekistan	13			
Croatia	1	Nicaragua	2	Venezuela	18			
Cuba	2	Niger	2	Yemen	1			
Ethiopia	14	Oman	3					
Fiji	1	Panama	5					
Only Brownfield (BF)								
Country	#BF	Country	#BF	Country	#BF	Country	#BF	
Aruba	1	Dominican Republic	1	Republic of the Congo	1			
Barbados	4	Eritrea	1	Sierra Leone	3			
British Virgin Islands	42	Isle of Man	1	Trinidad and Tobago	2			

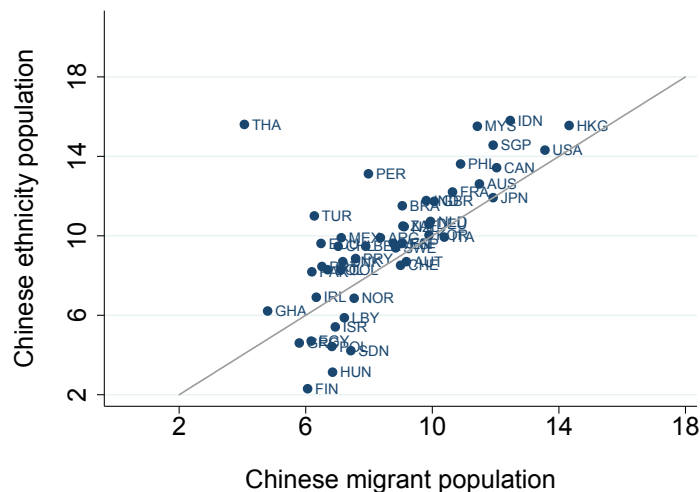
Appendix Table 2: Variable Definitions

<b>Variable</b>	<b>Definition</b>	<b>Source</b>
<i>#GF</i>	Number of Chinese outward greenfield investment projects	fDi intelligence
<i>GF</i>	Total investment value of Chinese outward greenfield projects	fDi intelligence
<i>#BF</i>	Number of cross-border M&A deals in host country with Chinese firms as the acquirers	SDC Platinum
<i>BF</i>	Total transaction value of cross-border M&A deals in host country with Chinese firms as the acquirers	SDC Platinum
<i>UN FDI</i>	Aggregate Chinese outward foreign direct investment flows	UNCTAD Bilateral FDI Statistics
<i>Migrant share</i>	Chinese emigrant population in host country as a share of host country's population	UN Global Migration Database & WDI
<i>(log) GDP</i>	(log) GDP in USD	WDI
<i>(log) GDP per capita</i>	(log) GDP per capita in USD	WDI
<i>Distance</i>	Geographic distance between China and host country	CEPII
<i>Language</i>	Dummy variable equal to 1 if at least 9% of population in host country speaks same language as China; zero otherwise	CEPII
<i>Legal System</i>	Dummy variable equal to 1 if host country has same legal origin as China (i.e., socialist)	La Porta et al. (1999)
<i>GDP growth</i>	Annual real GDP growth rate	WDI
<i>Financial Development</i>	Private credit by deposit money banks and other financial institutions divided GDP	Beck et al. (2000)
<i>Institutional Quality</i>	Sum of corruption, law and order, and bureaucratic quality indices (see Bekaert et al. (2004))	ICRG
<i>Trade Openness</i>	Sum of imports and exports divided by GDP	WDI
<i>Migrant share (40-year lag)</i>	Chinese migrant population in host country as a share of host country's population, from 1960-70	World Bank
<i>Dual citizenship</i>	Dummy variable equal to 1 if acquiring foreign citizenship does not lead to automatic loss of origin country's citizenship	MACIMIDE
<i>Immigration strictness</i>	Strictness of immigration policy associated with family reunification, asylum and refugees, and control of immigration	Helbling et al. (forthcoming)
<i>Immigrant hostility</i>	Share of population that do not want immigrants/foreign workers as neighbors	WVS
<i>Capital controls</i>	Dummy variable equal to 1 if capital controls are more strict than sample median	Gwartney et al. (2015)
<i>Labor regulations</i>	Dummy variable equal to 1 if hiring and firing regulations are less flexible than sample median	Gwartney et al. (2015)
<i>Administrative requirements</i>	Dummy variable equal to 1 if administrative requirements are more burdensome than sample median	Gwartney et al. (2015)

Notes: WDI = World Bank World Development Indicators. ICRG = International Country Risk Guide. WVS = World Values Survey.



Appendix Figure 1: Panels (a) and (b) plot histograms of the number of Chinese greenfield and brownfield investment projects over the whole sample from 2003-2014, respectively. The bin size for both panels is 10.



Appendix Figure 2: This figure plots the population of Chinese ethnicity against Chinese migrants for 49 countries in 1990.

Appendix Table 3: Summary Statistics

Variable	Mean	Median	Std Dev	Minimum	Maximum
# GF	4.85	2.00	8.71	1	80
GF (billion USD)	420	95.9	912	0.2	8954
# BF	3.50	2.00	4.83	1.00	36
BF (billion USD)	887	169	1530	0.02	8430
UN FDI (billion USD)	158	467	18.01	0.49	4808
Migrant share	0.002	0.0003	0.01	0	0.07
(log) GDP	24.2	24.1	2.44	18.6	30.5
(log) GDP per capita	8.68	8.71	1.53	5.25	11.6
(log) Distance	9.01	9.04	0.55	7.06	9.86
Language	0.02	0	0.13	0	1
Legal system	0.18	0	0.38	0	1
GDP growth	3.59	3.71	5.25	-62.1	104
Financial development	0.60	0.42	0.52	$9.9 \times 10^{-5}$	3.14
Institution quality	9.20	8.5	3.30	2.5	16
Trade openness	0.95	0.85	0.58	0.21	5.27
Migrant share (40-year lag)	0.002	0.0001	0.01	0	0.17
Dual citizenship	0.67	1	0.47	0	1
Immigration strictness	0.42	0.39	0.10	0.26	0.76
Immigrant hostility	0.23	0.19	0.16	0.02	0.67

Appendix Table 4: Robustness Checks with Host Country Fixed Effects

Dependent variable	Panel A: OLS			
	log (FDI + 1)			
	# GF	GF	# BF	BF
	(1)	(2)	(3)	(4)
Migrant share	27.893*	86.477	33.486*	147.633
	(14.817)	(64.545)	(19.623)	(99.030)
Observations	499	499	499	499
R-squared	0.766	0.706	0.688	0.532
Dependent variable	Panel B: Hausman-Taylor			
	log (FDI + 1)			
	# GF	GF	# BF	BF
	(1)	(2)	(3)	(4)
Migrant share	28.449**	92.062	34.194***	151.842**
	(11.410)	(72.443)	(9.286)	(72.075)
Observations	492	492	492	492

*Notes:* The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of green-field investment projects and likewise for M&A. The sample includes years 2003, 2005, 2010, and 2013. All regressions include the set of gravity equation control variables as described in the text and year fixed effects. All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.



Appendix Table 5: Robustness Check with (log) Migrant Stocks

	# GF (1)	GF (2)	# BF (3)	BF (4)
(log) Migrant stock	0.083*** (0.030)	0.265* (0.151)	0.129*** (0.036)	0.585*** (0.158)
(log) GDP	0.304*** (0.040)	1.616*** (0.179)	0.047 (0.030)	0.231 (0.164)
(log) GDP per capita	-0.251*** (0.056)	-1.210*** (0.253)	-0.062* (0.032)	-0.002 (0.213)
(log) Distance	0.070 (0.093)	-0.302 (0.428)	0.119 (0.087)	0.368 (0.485)
Language	0.311** (0.149)	2.120*** (0.750)	0.418 (0.304)	2.676 (1.913)
Legal system	0.177 (0.124)	0.817 (0.585)	0.093 (0.079)	0.111 (0.460)
GDP growth	0.003 (0.004)	0.045** (0.022)	0.004 (0.003)	0.033* (0.019)
Financial development	0.001 (0.001)	0.001 (0.004)	0.001 (0.001)	0.002 (0.005)
Institution quality	0.039 (0.025)	0.128 (0.114)	0.032 (0.020)	0.084 (0.110)
Trade openness	0.175** (0.071)	0.943*** (0.333)	-0.011 (0.083)	-0.122 (0.417)
Year fixed effects	Y	Y	Y	Y
Observations	1,016	1,016	1,016	1,016
R-squared	0.585	0.471	0.450	0.312

*Notes:* The dependent variables are Chinese outward FDI measured by (log) one plus the number and \$10,000 plus the value of greenfield investment projects ( $\#GF$  and  $GF$ ), and likewise for M&A ( $\#BF$  and  $BF$ ). All time-variant explanatory variables are lagged by one year. Standard errors in parentheses are clustered by host country. \*\*\*, \*\*, \* denote significance level at 1%, 5% and 10% respectively.