What drives Foreign Direct Investment: The role of language, geographical distance, information flows and technological similarity

Amadú Ly
BRU Research Fellow, ISCTE-IUL, Portugal

José Esperança
Full professor of Finance, ISCTE-IUL, Portugal

Nebojsa S. Davcik
BRU Research Fellow, ISCTE-IUL, Portugal

Abstract
This paper sheds new light on the impact of linguistic and technological similarities between countries on foreign direct investment (FDI), using an extended gravity model. The model includes technological commonality, as measured by the aggregate production of intellectual property, at the country level. An analysis of 71,309 pairs of FDI relationships, from 2000 to 2012, showed that language is positively associated with a high level of FDI. Technological differences do impede the flow of FDI between countries, and information flow is crucial for large flows of FDI. Information flow diminishes the negative impact of distance.

Key Words: Foreign direct investment, international business, language, distance, technology, information flow, gravity model

Acknowledgment

We are grateful to J.F. Hennart for his comments on previous versions of the manuscript and Jacinta Garcia from the World Tourism Organization (UNWTO) for providing the data on worldwide tourism flows. We are in debt for the detailed, constructive and valuable comments provided by the two anonymous referees and associate editor Piyush Sharma. This research was partially supported by Fundação para a Ciência e a Tecnologia, grant UID/GES/00315/2013. Amadú Ly would like to acknowledge the scholarship from Fundação para a Ciência e Tecnologia (Portugal).

Pre-print archiving statement: This is a pre-print of an article published in ***Journal of Business Research*****,** 2018, 88(7); 111-122. The definitive publisher-authenticated version is available online at: https://doi.org/10.1016/j.jbusres.2018.03.007

Suggested citation: Ly, A., Esperança, J., & Davcik, N. (2018). What drives Foreign Direct Investments (FDI): the role of language, geographical distance, information flow and technological similarity, ***Journal of Business Research***, *88*(7), 111-122.

What drives Foreign Direct Investment: The role of language, geographical distance, information flows and technological similarity

1. Introduction

Neoclassical theory predicts that capital will flow from rich to poor countries, and this will hold true until the returns from investments are equaled (Lucas, 1990). In fact, the bulk of foreign direct investment (FDI) take place in developed countries (UNCTAD, 1998), even though the highest returns can be obtained in developing countries (Pigato, 2000). Various and sometime competing explanations are offered in the debate about this phenomenon. Typically, the focus is on host countries' formal barriers, such as political risk, capital restrictions, taxes, legal and regulatory regime. Although it is intuitive that these factors affect FDI, the barriers to international investments have diminished considerably over the past few decades (Ahearne, Griever, & Warnock, 2004; Huberman, 2001). The sharp decline in transaction costs associated with FDI (e.g., similar legal background – UNCTAD, 1998), and diminished corporate tax rates (Devereux, Lockwood, & Redoano, 2008; Slemrod, 2004) have not diminished the skewed pattern of FDI toward developed countries. One possible explanation for this phenomenon is information asymmetry, because it is most acute in the international market (Doherty, 1999) and can hinder FDI. Information asymmetry can arise due to geographical distance (Coval & Moskowitz, 1999), or different legal and regulatory regime, or business practice between home and host countries (Ahearne et al., 2004). These differences are critical for FDI, because the high market entry costs are mainly costs of acquiring information regarding ways to conduct business in the host countries (Mata & Portugal, 2002). This situation is more severe in developing and emerging countries due to limited availability of public information (Kinoshita & Mody, 2001). Given that information is costly (Grossman & Stiglitz, 1980) and leads firms to equate unknown markets with high risks (Coval & Moskowitz, 1999), we expect it would skew investment preference toward countries that are familiar for investing firms. Access to information is made difficult by geographical distance between countries (Ivković & Weisbenner, 2005), thus leading to increased information asymmetry. Consequently, information costs are expected to increase in tandem with distance.

Therefore, we have identified factors which can impact distance and by extension minimize information asymmetry, namely, language, level of technological development and information flows. Theoretically, distance can be an incentive for FDI (Hirsch, 1976) or disincentive (Helpman, 1984), depending on the nature and purpose of the FDI. Language barriers contribute to information asymmetry among multinational corporations (MNCs), because they affect communication processes negatively (Kang & Kim, 2010). In addition, language differences between home and host countries means increased difficulty for MNCs in identifying business opportunities and negotiating agreements (Rauch & Trindade, 2002). Technology, particularly, information and communication technology (ICT), allows firms to circumvent barriers created by distance, enabling remote access to costumers and resources (Nachum & Zaheer, 2005), and reduces the costs of communications and coordination of operations (Mosakowski & Zaheer, 1999).

This research addresses the impact of information asymmetry on FDI in conjunction with the factors identified above. With the exceptions of Kinoshita and Mody (2001) and Loungani, Mody and Razin (2002), the majority of studies on information's effect on FDI are conceptual in nature (e.g., Goldstein & Razin, 2006). We sought to fill in this gap in the literature by using a dataset covering most of the world's economies and applying a different research methodology than that of previous studies. We used the Poisson Pseudo Maximum Likelihood (PPML) approach suggested by Silva and Tenreyro (2006), instead of the traditional ordinary least squares (OLS). Furthermore, we applied rarely used patent data as proxy for the level of technological development and tourism flow as proxy for information flow.

The present research further relied on the gravity model proposed by Pöyhönen (1963) and Tinbergen (1962). We used UNCTAD's outward bilateral FDI stock data and the Centre d' Études Prospectives et d'Informations Internationales's (CEPII) database for most of the gravity variables. We tested our hypotheses on panel data extending over 13 years (2000 - 2012).

The results have provided strong and positive evidence for the language effect on FDI, especially for high-income countries. The effects of geographic distance are strongly negative, for both high and low-income countries. We found that technological difference from high-income countries has a negative effect on FDI and a positive effect on FDI from low-income countries. Informational flow is positively associated with the level of FDI stock, even after accounting for possible tourism flow endogeneity. This result is not sensitive to source countries' income level or methodology applied.

This paper is organized as follows. In section two, we review the recent developments of studies on FDI determinants and desribe the research hypotheses. In section three, we describe the data used, and present the empirical model in section four. The results and robustness analysis are presented in sections five and six, respectively. Finally, section seven provides a summary and conclusions.

2. Literature Review

The impact of information asymmetry on capital flow has received considerable attention from both academics and policymakers (e.g., Portes, Rey & Oh, 2001; Portes & Rey, 2005; Tenzer, Terjesen, & Harzing, 2017). Most studies (e.g., Horstmann & Markusen, 1987; Goldstein & Razin, 2006) are conceptual or qualitative in nature, focusing on how information asymmetry can lead to one form of capital investment instead of another (e.g., FDI versus portfolio flow, and FDI versus licensing). The empirical studies, on the other hand, have been concerned with information asymmetry's impact on investors' behavior (e.g., Coval & Moskowitz 1999; Huberman, 2001; cf. Hejazi & Safarian, 2005). Rare empirical research exceptions in the FDI field are Kinoshita and Mody (2001) and Loungani et al. (2002). We present details on the conceptual nuances of FDI and information asymmetry (language, technology and distance) in International Business literature in Table 1.

Insert Table 1 here

2.1 Conceptual development of recent FDI research

2.1.1 Language in International Business (IB) literature

Contemporary language research in MNCs' literature has been predominantly qualitative, based on firm-level studies (e.g., O'Grady & Lane, 1996; Neeley et al., 2012). This research has highlighted the strong impact of language on IB, and the distinct effects of language and culture. In this study, by language we mean the same language type that is internationally standardized but with local differences in terms of dialect (a particular form of a language specific to a region), specific vocabulary or grammar tradition. For instance, the Portuguese language used in Portugal, Brazil, Cape Verde or Angola is the standard Portuguese in its essence. However, in each country the language has evolved revealing national differences in vocabulary, accent or grammar tradition due to external influences from local (native) culture and longstanding political independence from Portugal. Moreover, by language proximity we consider the language family that is a group of languages descending from a common language root. For instance, language commonalities between Portuguese and Spanish, or Spanish and Italian allow a certain ease of common communication between speakers without translation. In both cases, language is not the obstacle in communication or business interaction. This positioning is in line with recent research on language in IB literature by Tenzer et al. (2017) and Hejazi and Ma (2011).

The studies in IB literature have mostly focused on the post-FDI impact of language. Country level studies are critical to understand the impact of language on FDI. However, these studies have been rare in IB literature with the exceptions of Hejazi and Ma (2011); Oh, Selmier, and Lien (2011); and Selmier and Oh (2012). Although the literature has expanded our understanding of language influence on FDI, it has not, however, addressed the mechanisms through which language exerts its influence or the source countries' income level impact on language demand.

In this paper we explore the effect of information on FDI, taking into account the country of origin's effect on the pattern of FDI, because MNCs' attitudes toward risk and institutional factors are specific to a country's income level (Cuervo-Cazurra, 2012). Additionally, we used a language similarity variable approach as opposed to a binary variable one. Recognizing the fact that while two countries' languages might be different (e.g., Portuguese and Spanish), they can be understood and interchangeably used by their respective population, in such a way that they may significantly minimize transaction costs of investments (Selmier & Oh, 2012; cf. Zheng, 2014).

2.1.2 Distance in IB literature

There is an extensive literature on distance in the IB literature. Distance has long been perceived as a factor negatively affecting countries' relationship, since it is a source of friction between markets and produces greater transaction costs (Tesar & Werner, 1995). The studies on trade have confirmed these assertions (e.g., Rose, 2000; Frankel & Rose, 2002; Disdier & Head, 2008). However, IB literature (e.g., Hirsch, 1976; Horstmann & Markusen, 1987; Markusen & Venables, 1998) argue that if high export costs prevent an arm's length transaction, then setting up an operation in the form of a subsidiary could circumvent these problems. A competing view (e.g., Helpman, 1984) suggests that, if the purpose of FDI is to reduce costs and the relationship is mainly intra-firm, then the effect of distance will be to reduce FDI.

Initially, distance was used as a proxy for transport and communication costs. However, despite the continuous decline in transport and communication costs, the distance impact has not diminished. Some argue it has been rising over the years (Bénassy-Quéré, Coupet, & Mayer, 2007), causing researchers to question the overall effect of globalization. In most studies the distance estimate is strongly negative, even after controlling for factors such as colonial ties, common language, or membership in the same trading block (Ghemawat, 2001). Consequently, the attempts to explain the distance puzzle have shifted from transport and communication costs to information frictions. For instance, according to Rauch (1999) and Rauch and Trindade (2002) increased geographical distance means a higher cost of information acquiring, identifications costs (ex-ante), and higher information asymmetry among investors (Coval and Moskowitz, 1999). However, research on the implications for FDI has been scarce, with the exceptions of Kinoshita and Mody (2001) and Loungani et al. (2002).

2.1.3 Technology in IB literature

Technological differences between countries have long been regarded as a factor affecting the FDI flows. Differences in the level of development between countries limit information flows between firms and markets (Johanson & Vahlne, 1977). The interest among academics and policymakers in the link between technology and FDI is due to the belief that technology is a major driver of economic growth (e.g., Barrell & Pain, 1997; Borensztein, De Gregorio, & Lee, 1998; De Mello, 1999). However, the reluctance to abandon the traditional perspective of FDI flows from more advanced to developing economies has limited a broader analysis of contemporary FDI flows, as shown by Guillén and García-Canal (2009). Moreover, most FDI studies have ignored cultural variables such as language and its essential role in technology transfer.

In summary, FDI has attracted a large amount of research including the role of geographical distance and technology (see Table 1). However, to the best of our knowledge, no comprehensive study has addressed the joint impact of information asymmetry, distance, language similarity, information flow and the level of technological development on FDI. We believe this is an important gap in the literature, particularly because these variables individually are acknowledged as exerting significant influence on FDI patterns.

2.2 Hypotheses development

2.2.1 Language similarity and FDI

As previously noted, most research on the impact of language on IB has been qualitative firm-level studies (e.g., Neeley, Hinds, & Cramton, 2012). While these studies have enhanced our understanding of the complex nature of language in organizations, they are ex-post analysis on the effect of language on IB. Language is most critical ex-ante, in the initial phase of country selection and entry mode and the final phase of investment implementation. This is because language differences between home and host countries increase the difficulty MNCs experience in identifying market opportunities and negotiating business agreements (Rauch & Trindade, 2002).

Given that language barriers can negatively impact the levels of communication (Kang & Kim, 2010), this implies that language is a significant factor to consider in FDI decision-making. For the mobility of capital and interactions with diverse economic agents (e.g., public officials, suppliers and employees) require not only a close relationship and coordination, but also constant and high-quality information exchange.

A common language allows for easier communication and enhances trust (Lazear, 1999; Melitz, 2008; Rauch & Trindade, 2002) and minimizes information asymmetry between HQs and subsidiaries (Kang & Kim, 2010). Furthermore, language may reduce the impact of distance between countries and make distant locations attractive to potential investors (Hakanson & Ambos, 2010). Hence, the hypothesis:

H1: The stronger the language similarity, the larger will be the level of FDI.

2.2.2. Geographical distance and FDI

As noted above, the impact of distance on FDI remains a puzzle. Johanson and Wiedersheim-Paul (1975) assert that geographical distance is a component of the broader psychic distance between countries. Moreover, larger distance increases the costs of monitoring, coordinating and controlling of operations (Lerner, 1995), because of extra communication and transport costs (e.g., frequent visits, airfares, hotel stays and telephone calls – Petersen & Rajan, 1994). Furthermore, the fixed costs of setting up a plant abroad may be too high, rendering exporting the most efficient entry mode (Markusen & Venables, 1998). Hence, the hypothesis:

H2a: The higher the geographical distance, the smaller will be the level of FDI.

The steady decline in transportation and communication costs led to the belief that geographical distance should matter less (Bénassy-Quéré et al., 2007). However, the persistent large negative impact of distance challenges these assumptions (Ghemawat, 2001) and gave rise to a new explanation, information friction. The underlying assumption is that increased geographical distance increases the costs of information gathering (Lerner, 1995), limits information exchange and increases information asymmetry between investors (Coval & Moskowitz, 1999), undermining the feasibility of FDI. Distance in this respect is synonymous with difficult or costly information, and a large distance coefficient reflects barriers to information flows. Hence, the hypothesis:

H2b: The stronger the information flows, the smaller will be the impact of distance on FDI.

The distance impact is dependent on the specific characteristics of each pair of countries. Language similarity may play an important role in mitigating distance-based information asymmetry. This is because language similarity, in this particular setting, represents the quality of and access to information. It enhances the quality of information because it allows the understanding of nuances behind certain words and behaviors. Selmier and Ho (2012) argue that due to higher exposure to communication costs, language impact is higher for FDI than trade. Hence the hypothesis:

H2c: The higher the language similarity, the smaller will be the impact of distance on FDI.

2.3. Technological Similarity and FDI

As argued above, the conventional view on the impact of technology on FDI has limited the scope of research (Guillén & García-Canal, 2009). To understand the impact of technology on FDI we analyzed it in terms of firm and country level dimensions.

A MNC with very advanced ICT can gather, store and process a significant amount of information centrally and allocate this to dispersed units around the globe (Petersen & Rajan, 1994). In addition, ICT also allows firms to monitor operations effectively at arm's length, for instance, through profitability programs or automatic reporting systems (Petersen & Rajan, 1994). Moreover, an advanced ICT helps reduce information costs, accelerate the speed of information exchange and knowledge transfer between HQs and subsidiaries (Welch & Welch, 2008). In summary, an advanced ICT allows firms to circumvent barriers created by distance, enabling remote access to costumers and resources (Nachum & Zaheer, 2005). Hence, the hypothesis:

H3a: The higher the technological capability, the lower will be the impact of distance.

We further argue that the absorptive capability (e.g., individual language competence) (Welch & Welch, 2008), the stock of human capital availability and qualification and the general level of communication infrastructure in the host country are particularly important at a country level. These country specific characteristics attract FDI and help discriminate between competing locations (Dunning, 1980). Highly qualified human capital means that a country is capable of absorbing the most advanced technology available (Cohen & Levinthal, 1990), and MNCs' adjustment to foreign technology is quick and cost effective. An efficient communication infrastructure network plays a similar role, reducing the communication costs borne by MNCs. Moreover, if home and host country enjoy similar levels of technological development, *ceteris paribus*, the adaptation costs to set up a communication infrastructure for MNCs should also be lower, because the need to adjust the home country's technology to the host's is minimal. In contrast, two very different technological levels should be detrimental to FDI. Furthermore, a similar level of economic development allows MNCs to replicate their business models and exploit competitive advantage at a relatively low marginal cost. Hence, the hypothesis:

H3b: The larger the technological similarity, the higher will be the level of FDI.

The need for increased communication coupled with language diversity intensifies existing language barriers (Harzing & Feely, 2008), leading to slower and less efficient decision-making, as well as to power-distortion phenomena in HQ and subsidiary relationships (Harzing & Pudelko, 2013). In addition, reduced language competence leads to reduced absorptive capacity for subsidiaries during knowledge transfer (Welch & Welch, 2008), compromising both the MNCs' survival in highly competitive host markets and the host countries' welfare from FDI (Glass & Saggi, 1998). Hence, we argue that, given a minimum threshold of development and human capital qualifications, MNCs will prefer to invest in countries with similar language, despite different levels of technological development. Hence, the hypothesis:

H3c: The higher the language similarity, the smaller will be the impact of technological difference.

3. Data

We use panel data covering 13 years (2000-2012) for 224 countries and/or jurisdictions. The dataset consisted of 71,309 bilateral FDI stock observations of 649,376 country pairs. We use stock instead of flows, because flows are volatile and can significantly influence the results and question the interpretations (Bénassy-Quéré et al., 2007; Júlio et al., 2013). We deflated FDI stock data using US price deflator base year 2011, to obtain the "real" stock. Approximately 89% of the observations for the dependent variable is zero. About 71% of the FDI stock observations are from high-income countries and 29% are from low-income countries. For each hypothesis and year considered, the number of countries analyzed varies according to the data available. For instance, there is a lack of information in the developing countries' statistics for some variables. We used The World Bank gross national income (GNI) per capita criteria to determine the level of income of countries. For instance, GNI per capita larger than US$12,736 means that countries were categorized as high-income countries and low-income countries otherwise.

The gravity data was obtained from the CEPII's database. The language similarity variable came from Melitz and Toubal (2014), and tourism flow from the United Nations World Tourism Organization (UNWTO). Socioeconomic data (e.g., population, GDP, current US$), patent registration, level of schooling of the workforce) were taken from The World Bank World Development Indicators (WDI) (see Appendix A – Variables Descriptions and Sources).

4. Empirical method

4.1. Model

We use the gravity model suggested by Pöyhönen (1963) and Tinbergen (1962). The model proposes that objects (countries/economies) attract each other according to their mass/size (e.g., population, GDP), and the distance between countries reduces their attraction. This approach has been successfully used to explain bilateral FDI (e.g., Bevan & Estrin, 2004; Petroulas, 2007; Kleinert & Toubal, 2010).

We followed Silva and Tenreyro's (2006) nonlinear specification of the gravity model (PPML), because it deals with the zero observations in data and it is robust for different patterns of heteroscedasticity. In addition, the considerable number of zero observations in the FDI dataset renders the traditional ordinary least square (OLS) inadequate. The use of OLS in this specific case would lead to biased results or inconsistent estimates (Helpman et al., 2008) and would amplify the problem of outliers in OLS application. The literature (Silva & Tenreyro, 2006) shows that for data sets with characteristics as in this study, PPML provides a better fit and more robust estimates than other methods, including Tobit. We use an extended gravity model as in Kleinert and Toubal (2010) and Loungani et al. (2002). We intentionally do not consider using fixed effect models, because these do not allow for estimations of time invariant regressors such as geographical distance and language. We specified the following regression equation:

$$lnFDI_{ijt} = \beta_0 + \beta_1 Distance_{ij} + \beta_2 Language_{ij} + \beta_3 Tech_{ijt} + \beta_4 Infl_{ijt} + \beta_5 Control_{ijt} + \varepsilon_{ijt} \quad (1)$$

Where *i* and *j* are countries, *t* is time, and FDI_{ijt} is the deflated outward bilateral FDI stock between countries *i* and *j* at time *t*. $Distance_{ij}$ corresponds to a vector of a variable representing distance (e.g., $Distance_{ij}$ = distance between countries *i* and *j*; $Cont_{ij}$ = contiguity, a binary variable that represents neighboring countries). $Language_{ij}$ represents language similarity variable between countries *i* and *j* (*language_simasjp*). Tech is a vector of variables representing the level of technology of countries *i* and *j* (e.g., $SimilarPat_{ij}$). The $Infl_{ijt}$ variable represents the information flow between countries *i* and *j* in period *t* (e.g., *Std_Tourism*). Control is a vector of control variables as per the current literature on FDI determinants (e.g., GDP, GDP per Capita, GDP growth rate, inflation rate, workforce education, legal origin, religion, currency union, and exchange rate) between countries *i* and *j* in period *t*. $\mathcal{E}_{ijt}$ represents the error term.

4.2. Variables

The dependent variable is the deflated outward bilateral FDI stock between countries *i* and *j* (FDI_{ijt}). Following Silva and Tenreyro (2006), the variable was computed in levels.

Deflated GDP (GDP_{ijt}) between countries *i* and *j* in period *t* can be regarded as a proxy for supply and demand forces (Rose, 2000), so we expected it to be positively related to FDI.

Geographical distance between countries *i* and *j* ($Dist_{ij}$) was measured by the distance, in kilometers, between the capital cities of the countries *i* and *j*.

Language similarity ($language\_simasjp_{ij}$) between countries *i* and *j*, was formulated as a continuous index, with zero as lower bound and one as the upper bound (zero = no similarity; one = same language).

Patent (*pat*) represented the number of patents applications filed by residents of a given country at the national patent office. We computed the level of technological similarity (*SimilarPat*) by adjusting Egger and Pfaffermayr (2004)'s economic similarity index:

$$SimilarPat = ln\left(1 - \left(\frac{Pt_{orig}}{Pt_{orig} + Pt_{dest}}\right)^2 - \left(\frac{Pt_{dest}}{Pt_{orig} + Pt_{dest}}\right)^2\right) \quad (2)$$

Tourism was defined as a flow measure of foreign visitors between countries *i* and *j*. This variable was normalized (*St_Tourism*) as in Portes et al. (2001), to remove the size effect:

$$ln\left(\frac{Tourism\_orig+Tourism\_dest}{\sqrt{(Defgdp_{orig}*Defgdp_{dest})}}\right) \quad (3)$$

5. Results and discussion

The estimation results, using PPML, are presented in Table 2. We also split the samples of FDI into high and low-income countries using the World Bank's GNI per capita criteria—as described in section 3 and shown in Tables 3 and 4.

Insert Tables 2, 3, and 4.

All the variables present the expected signs, and are economically and statistically significant. The market size variables, GDPs of both home and host countries are positive and economically significant across all samples.

As to the variables of interest, the language similarity variable (*language_simasjp*) is positive and highly significant (1.288) (Model 2, Table 2), meaning that it increases the level of FDI stock 3.63 times (exp(1.288)). This result emphasizes the strong influence of language on FDI decision. This estimate is higher than that of Loungani et al. (2002) (0.749), but lower than that of Bénassy-Quéré et al. (2007) (1.77). For the high-income sample the language variable is positive and statistically significant (Model 2, Table 3). This finding suggests that priority must be given to workforce language qualification to attract FDI from these countries. In contrast, the evidence is not as strong for low-income countries, because the result is not statistically significant (Model 2, Table 4).

The geographical variables (distance, contiguity, and landlocked) (2a) show the expected sign. The coefficient for distance (*LnDist*) is negative and highly significant (Model 1, Table 2), suggesting that a 1% increase in distance between a given pair of countries represents a reduction in FDI in the order of 0.846%. This coefficient estimate is below that of Loungani et al. (2002) (-1.199), but higher than that of Júlio et al. (2013) (-0.637) and Bénassy-Quéré et al. (2007) (-0.53). We replicated the same results for the high and low-income samples (Model 1, Tables 3 and 4). We found a lower distance coefficient for low-income countries compared to high-income countries, suggesting MNCs of the former are less sensitive to distance than those from the latter. Perhaps because of their limited market opportunities, low-income countries' MNCs go larger distances to secure business deals (Cuervo-Cazurra & Genc, 2008).

To test for the information costs underlying the distance effect, that is the effect of information on distance (2b), we use the bilateral tourism flow (*Std_Tourism*) as in Portes et al. (2001). The result is positive and statistically significant (Model 3, Table 2). In addition, we found that information flow reduces the distance coefficient to 0.361. This result is slightly higher than the theoretical upper margin suggested by Blonigen et al. (2002). We obtained similar results across the other samples (Model 3, Tables 5 and 6). These results imply that increasing information flow may help in attracting FDI. Interestingly, for low-income countries the distance becomes positive, but not statistically significant. This finding is in line with market seeking objectives as suggested by Loungani et al. (2002).

To test the effect of language on distance (2c), we added the language similarity variable to Model 1 (Table 2), and the result is a decline in the distance coefficient. The distance coefficient shifts from 0.846 (Model 1, Table 2) to 0.767 (Model 2, Table 2). This result favors the notion of information costs of distance. In addition, we cross-check the result by interacting the distance and language, under the assumption that language compensates for distance between countries (Hakanson & Ambos, 2010). We plotted the marginal effects of all the interacted variables to check the behaviors and signs. This approach is in line with Ai and Norton (2003) and Greene (2010). The interaction variable (*Distlanguage_simasjp*) is positive and statistically significant for the overall and high-income sample (Model 4, Tables 2 and 3) but not statistically significant for the low-income sample (Model 4, Table 4). These results reinforce our hypothesis that language is a significant factor that minimizes distance, particularly for high-income countries.

To test for the effect of technological level on FDI (3a), we first added the variables representing each country's level of technology. The results suggest home country level of technology is critical for FDI. In addition, these variables diminish the distance coefficient to 0.721 (Model 5, Table 2). To cross-check the results, we also tested the interaction effects of the home and host countries' technology with distance. The interaction variable (*DistMultPt*) is positive and statistically significant for the overall and high-income samples (Model 7, Tables 2 and 3), as well as negative and not statistically significant for low-income countries (Model 7, Table 4). This would suggest high-income countries' MNCs use technology to minimize the distance to host countries.

To test the effect of technological similarity on FDI (3b), we used the level of technological similarity (*SimilarPat*) expressed by Equation 2. The result is not statistically significant for the overall sample (Model 6, Table 2), not statistically significant for the high-income sample (Model 6, Table 3), and negative statistically significant for the low-income sample (Model 6, Table 4). The overall results are, therefore, inconclusive. This finding suggests that investments preference lean toward countries of different levels of development than the home country. However, for the low-income sample there is a clear preference for technologically dissimilar countries when investing abroad. This is in line with the prevailing logic about the behavior of the low-income countries' MNCs (Cuervo-Cazurra, 2012; Guillén & García-Canal, 2009). For high-income countries the evidence, although not strong, points to investment in countries with similar factor endowment.

Finally, to test the effect of language on technological difference (3c), we used the interaction of the language and technological difference variables. The rationale is that although technological difference hinders FDI, it should be less important in the presence of language similarities between countries. The results for the interaction variable (*LogDifPtlanguage_simasjp*) are positive and statistically significant for the overall and high-income samples (Model 8, Tables 2 and 3), and positive but not statistically significant for the low-income sample (Model 8, Table 4). These findings suggest that language can act as a bridge between countries at different stages of development, particularly for high-income countries.

6. Robustness check

We adopted other proxy measures to cross-check the results in Table 2. We used different proxy measure for information flow (e.g., existence of stock exchange, IMF loan, Multilateral Investment Guarantee Agency), and level of technological level (e.g., fixed broadband internet subscriber). The new estimations confirmed the results in Table 2 (results are available upon request). Additionally, we also tested for the endogeneity of tourism flow by instrumenting the intensity of fixed telephone subscriber. Once we account for this, both the distance and language similarity variables are statistically insignificant, reinforcing the idea of information costs being inherent to distance and language. We also tested for the endogeneity of technology by instrumenting the intensity of energy consumption. Once we account for this, both distance and language diminish considerably and in the case of FDI from low-income countries the distance variable becomes insignificant. Moreover, we also tested for institutional variables (e.g., bureaucratic inefficiencies, corporate tax rate, rule of law, and control of corruption). The results showed that differences in institutional variables tend to make countries more distant and language similarities less effective. Furthermore, we carried out OLS estimation with a robust standard error for all models in Table 4, to confirm the robustness of the results above. These analyses have shown that our findings are not driven by the particular method or dependent variable. The model specification is in line with methodological standards in the literature.

$ln(\alpha+FDI_{ijt}) = \beta_0 + \beta_1 Distance_{ij} + \beta_2 Language_{ij} + \beta_3 Tech_{ijt} + \beta_4 Infl_{ijt} + \beta_5 Control_{ijt} + \varepsilon_{ijt}$ (4), where α is greater than zero and less or equal to 1.

All variables have the same sign as PPML estimates (Appendix B – Results of the OLS regression), except for 4 variables, i.e., the difference in exchange rate, growth rate, information flow, technological similarity. These differences lie in the magnitude of the coefficients in which PPML estimates show better fit.

7. Discussion

7.1. Theoretical contribution

This study sheds new light on the link between information asymmetry and FDI (i) providing empirical evidence in a debate that has been mostly conceptual or qualitative in nature and (ii) applying a different methodological approach (a gravity model using PPML) that provides better and more robust estimates than classical OLS techniques. Distance is a significant obstacle to FDI, as it makes communication and interactions between countries difficult. However, we found that high-income countries' MNCs are more sensitive to distance than their low-income counterparts. Perhaps, this is due to the fact that low-income countries' MNCs must travel longer distances to secure business deals (Cuervo-Cazurra, 2012). Interestingly, we confirmed that the distance impact can be mitigated by ICT and language similarity.

We found that information flow is a significant factor in attracting FDI, irrespective of the source countries' income level. In addition, information flows also contribute to minimize the distance between countries. More importantly, these flows appear to signal the market searching nature of FDI from low-income countries' MNCs and the quest for cost reduction for high-income countries' MNCs. This result is robust to the consideration of possible endogeneity of tourism flow. In fact, the large impact of language and distance dissipates once we take into account the possible endogeneity effect. These findings reinforce the idea of distance as a proxy for information asymmetry and that language impact is due mainly to information costs. This result is in line with Loungani et al. (2002).

We also found that language similarity is an important factor to attract FDI. However, we observed distinct patterns of behavior depending on source countries' income level. For high-income countries, language is far more significant to attract FDI then for low-income countries. These findings are significant because previous studies on the effect of language (e.g., Selmier & Oh, 2012) have not broken down the results by income level and they have assumed the impact is uniform across source countries. In addition, understanding FDI dynamics is important due to the changes in international institutional setup and business environment. It is commonly accepted in IB literature that language proximity fosters trade between two nations. However, the recent developments with a country such as the UK exiting the EU shows that trade flows between political and economic blocks (such as the EU & USA, or NAFTA & EU) have different patterns[1] and suggest counter intuitive results. During the first period of the Brexit political waffle, the main argument was that the UK does not need the EU as they can trade with the USA, Canada and Australia. Unfortunately, that was an ill argument, because all current commercial arrangements that the UK has are under the umbrella of the EU treaties with other countries or trade blocks. Once the UK leaves the EU (which is expected in 2019), it must start new commercial negotiations and those outcomes will not necessarily be better for the UK than those under the EU umbrella. Ultimately, the USA decided to give priority to the EU rather than to the UK in further trade deregulation, in spite of the language proximity, high-income country performance and political closeness between the USA and UK.

We found that technological differences between countries to hinder FDI. However, these results are also specific to the source countries' income level. Low-income countries' MNCs prefer dissimilar economies for investment as suggested in the literature in search for technology and markets (Cuervo-Cazurra & Genc, 2008). Interestingly, ICT minimizes the distance between countries more clearly for high-income countries. Furthermore, language similarity minimizes the effect of the technological difference between countries.

7.2. Practical implications for managers and policymakers

The study shows a different pattern of FDI depending on the source countries' income level. We found that high-income countries' MNCs take a more conservative approach to investments than their low-income counterparts, because the former prefer a shorter distance, similar language, and the

[1] We are grateful to the Reviewer #1 for pointing out this theoretical implication.

same level of technological development. Conversely, low-income countries' MNCs are less stringent on their expectations regarding the surrounding environment. This result implies that the motivations for FDI by these two types of MNCs are different as suggested by anecdotal evidence presented in studies such as Mathews (2006) or Guillen and Garcia-Canal (2009). These findings are significant for policymakers and academics interested in the geography of FDI, because they allow for a better understanding of the nature of market forces in different countries.

The results of this paper have important implications for countries seeking to attract FDI and companies searching business opportunities in foreign markets. The findings suggest that improving information and access to information by host countries' government agencies is vital to attract FDI. For instance, low-income countries could adopt international organizations' benchmarks and reporting standards (e.g., UNCTAD, 2006), making data more easily accessible to potential investors. This study's results are also encouraging for countries distant from the main financial and decision centers, as the findings point to forms of mitigating distance barriers. In addition, this research provides the grounds for customizing policy for specific kinds of FDI, for instance, improving language qualification of the workforce to attract high-income countries' FDI.

7.3. Limitations and paths for future research

This study has several limitations that should inspire further research. We have not tested the effect of a *lingua franca* (i.e., English) on the overall language similarity result. However, the results obtained are significantly robust, confirming the impact of language similarity on FDI.

The second limitation arises from the absence of firm-level data. A more subtle analysis should focus on the impact of industry and firm specific variables such as concentration, patents, size, age or governance. The different patterns of high-income countries' MNCs and the "new" MNCs from other regions suggest that different capabilities are emerging alongside, supplementary to the traditional technological, financial and managerial advantages that compensate the liability of foreignness.

Finally, the objective limitation of the study is the quality of data that is available. Notwithstanding our effort, better formatted and more detail subsectors would have improved the current study. The increasing availability of large-scale data on FDI, advances in big data analytics and software shall likely foster future research on FDI and variables that affect the IB environment at the firm and country levels in much greater detail.

Our results show that, for FDI, distance matters, but language similarity, information flows and technological similarity have potentially moderating effects. This is particularly true for new MNCs from lower income countries. Firm based observation focusing on host country language and language capabilities, organizational structure and technological skills would enhance our knowledge of the impact of these variables for different industries and firms. This stream of research may provide a rich field for testing the current explanations and boundaries of foreign expansion by MNCs.

References

Ahearne, A. G., Griever, W. L., & Warnock, F. E. (2004). Information costs and home bias: an analysis of US holdings of foreign equities. *Journal of International Economics, 62*(2), 313-336.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters, 80*(1), 123-129.

Álvarez, I., & Marin, R. (2013). FDI and Technology as Levering Factors of Competitiveness in Developing Countries. *Journal of International Management, 19*(3), 232-246.

Barrell, R., & Pain, N. (1997). Foreign direct investment, technological change, and economic growth within Europe. *The Economic Journal, 107*(445), 1770-1786.

Bénassy-Quéré, A., Coupet, M., & Mayer, T. (2007). Institutional determinants of foreign direct investment. *The World Economy, 30*(5), 764-782.

Bevan, A. A., & Estrin, S. (2004). The determinants of foreign direct investment into European transition economies. *Journal of Comparative Economics, 32*(4), 775-787.

Blonigen, B. A., Davies, R. B., & Head, K. (2002). Estimating the knowledge-capital model of the multinational enterprise: Comment: National Bureau of Economic Research.

Borensztein, E., De Gregorio, J., & Lee, J.-W. (1998). How does foreign direct investment affect economic growth? *Journal of International Economics, 45*(1), 115-135.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 128-152.

Coval, J. D., & Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance, 54*(6), 2045-2073.

Cuervo-Cazurra, A., & Genc, M. (2008). Transforming disadvantages into advantages: developing-country MNEs in the least developed countries. *Journal of International Business Studies, 39*(6), 957-979.

Cuervo-Cazurra, A. (2012). Extending theory by analyzing developing country multinational companies: Solving the Goldilocks debate. *Global Strategy Journal, 2*(3), 153-167.

Davidson, W. H., & Mcfetridge, D. G. (1985). Key Characteristics in the Choice of International Technology-Transfer Mode. *Journal of International Business Studies, 16*(2), 5-21.

De Mello, L. R. (1999). Foreign direct investment-led growth: evidence from time series and panel data. *Oxford Economic Papers, 51*(1), 133-151.

Devereux, M. P., Lockwood, B., & Redoano, M. (2008). Do countries compete over corporate tax rates? *Journal of Public Economics, 92*(5), 1210-1235.

Disdier, A.-C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics, 90*(1), 37-48.

Doherty, A. M. (1999). Explaining international retailers' market entry mode strategy: internalization theory, agency theory and the importance of information asymmetry. *The International Review of Retail, Distribution and Consumer Research, 9*(4), 379-402.

Dunning, J. H. (1980). Toward an Eclectic Theory of International Production: Some Empirical Tests. *Journal of International Business Studies*, 11(1), 9-31.

Eaton, J., & Tamura, A. (1994). Bilateralism and regionalism in Japanese and US trade and direct foreign investment patterns. *Journal of the Japanese and International Economies, 8*(4), 478-510.

Egger, P., & Pfaffermayr, M. (2004). The impact of bilateral investment treaties on foreign direct investment. *Journal of Comparative Economics, 32*(4), 788-804.

Frankel, J., & Rose, A. (2002). An estimate of the effect of common currencies on trade and income. *The Quarterly Journal of Economics, 117*(2), 437-466.

Ghemawat, P. (2001). Distance still matters - The hard reality of global expansion. *Harvard Business Review, 79*(8), 137-147.

Glass, A. J., & Saggi, K. (1998). International technology transfer and the technology gap. *Journal of Development Economics, 55*(2), 369-398.

Goldstein, I., & Razin, A. (2006). An information-based trade-off between foreign direct investment and foreign portfolio investment. *Journal of International Economics, 70*(1), 271-295.

Greene, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters, 107*(2), 291-296.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review, 70*(3), 393-408.

Guillén, M. F., & García-Canal, E. (2009). The American model of the multinational firm and the “new” multinationals from emerging economies. *The Academy of Management Perspectives, 23*(2), 23-35.

Hakanson, L., & Ambos, B. (2010). The antecedents of psychic distance. *Journal of International Management, 16*(3), 195-210.

Harzing, A.-W., & Feely, A. J. (2008). The language barrier and its implications for HQ-subsidiary relationships. *Cross Cultural Management: An International Journal, 15*(1), 49-61.

Harzing, A.-W., & Pudelko, M. (2013). Language competencies, policies and practices in multinational corporations: A comprehensive review and comparison of Anglophone, Asian, Continental European and Nordic MNCs. *Journal of World Business, 48*(1), 87-97.

Hejazi, W., & Ma, J. (2011). Gravity, the English language and international business. *Multinational Business Review, 19*(2), 152-167.

Hejazi, W., & Safarian, A. E. (2005). NAFTA effects and the level of development, *Journal of Business Research*, *58*(12), 1741-1749.

Helpman, E. (1984). A simple theory of international trade with multinational corporations. *The Journal of Political Economy, 92*(3), 451-471.

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics, 123*(2), 441-487.

Hirsch, S. (1976). An international trade and investment theory of the firm. *Oxford Economic Papers,* 258-270.

Horstmann, I., & Markusen, J. R. (1987). Licensing Versus Direct-Investment - a Model of Internalization by the Multinational-Enterprise. *Canadian Journal of Economics-Revue Canadienne D Economique, 20*(3), 464-481.

Huberman, G. (2001). Familiarity breeds investment. *Review of Financial Studies, 14*(3), 659-680.

Ivković, Z., & Weisbenner, S. (2005). Local does as local is: Information content of the geography of individual investors' common stock investments. *The Journal of Finance, 60*(1), 267-306.

Johanson, J., & Vahlne, J.-E. (1977). The internationalization process of the firm—a model of knowledge development and increasing foreign market commitments. *Journal of International Business Studies, 8*(1), 23-32.

Johanson, J., & Wiedersheim-Paul, F. (1975). The internationalization of the firm—four Swedish cases 1. *Journal of Management Studies, 12*(3), 305-323.

Júlio, P., Pinheiro-Alves, R., & Tavares, J. (2013). Foreign direct investment and institutional reform: evidence and an application to Portugal. *Portuguese Economic Journal*, *12*(3), 215-250.

Kang, J.-K., & Kim, J.-M. (2010). Do foreign investors exhibit a corporate governance disadvantage? An information asymmetry perspective. *Journal of International Business Studies, 41*(8), 1415-1438.

Kinoshita, Y., & Mody, A. (2001). Private information for foreign investment in emerging economies. *Canadian Journal of Economics/Revue canadienne d'économique, 34*(2), 448-464.

Kleinert, J., & Toubal, F. (2010). Gravity for FDI. *Review of International Economics, 18*(1), 1-13.

Lazear, E. (1999). Language and culture. *Journal of Political Economy, 107*(6), S95-S126.

Lerner, J. (1995). Venture capitalists and the oversight of private firms. *The Journal of Finance, 50*(1), 301-318.

López-Duarte, C., & Vidal-Suárez, M. M. (2010). External uncertainty and entry mode choice: Cultural distance, political risk and language diversity. *International Business Review, 19*(6), 575-588.

Loungani, P., Mody, A., & Razin, A. (2002). The global disconnect: the role of transactional distance and scale economies in gravity equations. *Scottish Journal of Political Economy, 49*(5), 526-543.

Lucas, R. E. (1990). Why doesn't capital flow from rich to poor countries? *The American Economic Review*, 92-96.

Markusen, J. R., & Venables, A. J. (1998). Multinational firms and the new trade theory. *Journal of International Economics, 46*(2), 183-203.

Mata, J., & Portugal, P. (2002). The survival of new domestic and foreign-owned firms. *Strategic Management Journal, 23*(4), 323-343.

Mathews, J. (2006). Dragon Multinationals. *Asia Pacific Journal of Management*, *23*, 5-27.

Melitz, J. (2008). Language and foreign trade. *European Economic Review, 52*(4), 667-699.

Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. Journal of International Economics, 93 (2), 351-363

Mosakowski, E., & Zaheer, S. (1999). The global configuration of a speculative trading operation: an empirical study of foreign exchange trading. *Organization Science, 10*(4), 401-423.

Nachum, L., & Zaheer, S. (2005). The persistence of distance? The impact of technology on MNE motivations for foreign investment. *Strategic Management Journal, 26*(8), 747-767.

Neeley, T. B., Hinds, P. J., & Cramton, C. D. (2012). The (un) hidden turmoil of language in global collaboration. *Organizational Dynamics, 41*(3), 236-244.

O'Grady, S., & Lane, W. H. (1996). The Psychic Distance Paradox, *Journal of International Business Studies, 27*(2), 309-333.

Oh, C. H., Selmier, T., & Lien, D. (2011). International trade, foreign direct investment, and transaction costs in languages. *The Journal of Socio-Economics, 40*(6), 732-735.

Petersen, M. A., & Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *The Journal of Finance, 49*(1), 3-37.

Petroulas, P. (2007). The effect of the euro on foreign direct investment. *European Economic Review, 51*(6), 1468-1491.

Pigato, M. (2000). *Foreign direct investment in Africa: old tales and new evidence*: World Bank.

Portes, R., & Rey, H. (2005). The determinants of cross-border equity flows. *Journal of International Economics, 65*(2), 269-296.

Portes, R., Rey, H., & Oh, Y. (2001). Information and capital flows: The determinants of transactions in financial assets. *European Economic Review, 45*(4), 783-796.

Pöyhönen, P. (1963). A tentative model for the volume of trade between countries. *Weltwirtschaftliches Archiv*, 93-100.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics, 48*(1), 7-35.

Rauch, J. E., & Trindade, V. (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics, 84*(1), 116-130.

Rose, A. K. (2000). One money, one market: the effect of common currencies on trade. *Economic Policy, 15*(30), 7-46.

Selmier, W. T., & Oh, C. H. (2012). International business complexity and the internationalization of languages. *Business Horizons, 55*(2), 189-200.

Selmier, W. T., & Oh, C. H. (2012). The Power of Major Trade Languages in Trade and Foreign Direct Investment. *Review of International Political Economy, 20*(3), 486-514.

Silva, J. S., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics, 88*(4), 641-658.

Slemrod, J. (2004). Are corporate tax rates, or countries, converging? *Journal of Public Economics, 88*(6), 1169-1186.

Smarzynska, B. K. (2002). *The composition of foreign direct investment and protection of intellectual property rights: evidence from transition economies* (Vol. 2786): World Bank Publications.

Tenzer, H., Terjesen, S., & Harzing, A.-W. (2017). Language in International Business: A Review and Agenda for Future Research, *Management International Review*, *57*, 815-854.

Tesar, L. L., & Werner, I. M. (1995). Home bias and high turnover. *Journal of International Money and Finance, 14*(4), 467-492.

Tinbergen, J. (1962). An analysis of world trade flows. In *Shaping the world economy*. Tinbergen, J. (ed.). New York NY: Twentieth Century Fund

UNCTAD (1998). *World Investment Report 1998: Trends and Determinants*. United Nations Conference on Trade and Development, United Nations Publications, New York and Geneva.

UNCTAD (2006). *FDI from Developing and Transitioning Economies: Implications for development*. *World Investment Report 2006*. United Nations Conference on Trade and Development, United Nations Publications, New York and Geneve.

Welch, D. E., & Welch, L. S. (2008). The importance of language in international knowledge transfer. *Management International Review, 48*(3), 339-360.

Zheng, P. (2014). Antecedents to international student inflows to UK higher education: A comparative analysis, *Journal of Business Research*, *67*(2), 136-143.

Table 1 – Theoretical Summary

Authors	FDI	Language	Geographical Distance	Technology	Information	Portfolio	Method
Álvarez and Marin (2013)	+			+			GMM
Barrell and Pain (1997)	+			+			OLS
Bénassy-Quéré et al. (2007)	+	+	+				OLS
Blomström and Sjöholm (1999)	+			+			OLS
Blonigen et al. (1997)	+						Negative binomial
Borensztein et al. (1998)	+			+			OLS
Davidson and Mcfetridge (1985)	+	+	+	+			Logit
Eaton and Tamura (1994)	+						Modified Tobit
Goldstein and Razin (2006)	+				+	+	Qualitative/Theoretical
Hejazi and Ma (2011)	+	+	+				OLS
Hortsmann and Markusen (1987)	+				+		Qualitative
Kinoshita and Mody (2001)	+		+		+		Multinomial logit
López-Duarte and Vidal-Suárez (2010)	+	+					Logistic regression (binomial)
Loungani et al. (2002)	+	+	+		+		Tobit
Nachum and Zaheer (2005)	+		+	+			OLS
Neeley et al. (2012)		+					Qualitative
O'Grady and Lane (1996)		+					Qualitative
Oh et al. (2011)	+	+	+				OLS
Portes and Rey (2005)		+	+	+	+	+	OLS
Portes et al. (2001)		+	+	+	+	+	OLS
Rauch and Trindade (2002)		+	+				Modified Tobit
Selmier and Oh (2013)	+	+	+		+		OLS
Smarzynska (2002)	+			+			Probit/ multinomial logit/OLS

Note: OLS = ordinary least square; GMM = generalized method of moments

Table 2: Results of PPML Regression

	Control Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Common Colonizer	-0.551*** (0.205)	0.195 (0.171)	0.864*** (0.168)	2.218*** (0.479)	0.894*** (0.169)	0.880*** (0.164)	0.767*** (0.156)	0.859*** (0.163)	0.731*** (0.160)	0.859*** (0.169)
Colony After 1945	-0.298 (0.184)	-0.285** (0.141)	0.235 (0.153)	2.505*** (0.363)	0.310** (0.158)	0.251 (0.159)	0.175 (0.180)	0.245 (0.157)	0.200 (0.153)	0.234 (0.153)
Same Country	0.325** (0.154)	0.498*** (0.133)	0.564*** (0.146)	1.045*** (0.229)	0.390*** (0.146)	0.658*** (0.148)	0.596*** (0.146)	0.667*** (0.151)	0.564*** (0.147)	0.558*** (0.147)
legorig_dumm	0.336*** (0.0899)	0.562*** (0.0689)	0.250*** (0.0802)	-0.349*** (0.112)	0.228*** (0.0802)	0.506*** (0.0834)	0.282*** (0.0791)	0.481*** (0.0868)	0.277*** (0.0807)	0.249*** (0.0802)
rlg_dumm	0.128 (0.0786)	0.122** (0.0623)	0.116* (0.0622)	-0.0453 (0.0826)	0.0674 (0.0627)	0.0621 (0.0642)	0.111* (0.0640)	0.0685 (0.0647)	0.103 (0.0631)	0.116* (0.0623)
currunion_dum	1.088*** (0.0839)	0.110 (0.0741)	0.0448 (0.0701)	0.247** (0.120)	0.0291 (0.0683)	0.0506 (0.0701)	0.0734 (0.0730)	0.0497 (0.0691)	0.0739 (0.0744)	0.0476 (0.0703)
rta	0.0925 (0.118)	-0.925*** (0.131)	-0.491*** (0.137)	-0.298 (0.311)	-0.521*** (0.138)	-0.460*** (0.142)	-0.533*** (0.144)	-0.470*** (0.147)	-0.536*** (0.143)	-0.488*** (0.136)
LogDifExchrt	-0.381** (0.188)	-0.449 (0.295)	-0.396 (0.252)	-2.673*** (0.548)	-0.419* (0.241)	-0.386 (0.238)	-0.428* (0.250)	-0.491** (0.234)	-0.429* (0.250)	-0.391 (0.250)
DifGrwthRt	3.176*** (0.870)	0.428 (1.002)	1.242 (1.104)	1.343 (1.408)	1.162 (1.108)	1.162 (1.116)	0.919 (1.116)	0.892 (1.096)	1.140 (1.139)	1.246 (1.101)
LogDifInfl	0.0171 (0.0520)	-0.0245 (0.0321)	-0.00385 (0.0332)	-0.0530 (0.0414)	-0.00594 (0.0328)	-0.0225 (0.0348)	-0.0162 (0.0349)	-0.0203 (0.0353)	-0.0116 (0.0354)	-0.00375 (0.0331)
DifLbFcSec	-0.192*** (0.0458)	0.0155 (0.0634)	-0.160*** (0.0558)	-0.381** (0.167)	-0.154*** (0.0563)	-0.424*** (0.0858)	-0.139** (0.0565)	-0.139** (0.0626)	-0.196*** (0.0591)	-0.160*** (0.0560)
DifLbFcTer	-0.0628 (0.0430)	0.0107 (0.0596)	0.0398 (0.0565)	-0.186*** (0.0712)	0.0340 (0.0564)	-0.00140 (0.0618)	0.0546 (0.0611)	0.0651 (0.0612)	0.0728 (0.0613)	0.0389 (0.0565)
GDP Orig		0.850*** (0.0320)	0.908*** (0.0291)	0.875*** (0.0721)	0.894*** (0.0291)	0.573*** (0.0385)	0.897*** (0.0273)	0.848*** (0.0282)	0.872*** (0.0287)	0.906*** (0.0290)
GDP CapOrig		0.00423 (0.0411)	-0.0267 (0.0436)	0.385* (0.198)	-0.0117 (0.0438)	0.242*** (0.0528)	-0.0304 (0.0439)	0.153*** (0.0585)	-0.00242 (0.0435)	-0.0256 (0.0436)
GDP Dest		0.477*** (0.0277)	0.587*** (0.0333)	0.623*** (0.0477)	0.576*** (0.0337)	0.593*** (0.0349)	0.591*** (0.0299)	0.545*** (0.0307)	0.606*** (0.0356)	0.585*** (0.0333)
GDP CapDest		0.499*** (0.0434)	0.216*** (0.0474)	-0.443*** (0.0923)	0.224*** (0.0470)	0.300*** (0.0455)	0.198*** (0.0496)	0.330*** (0.0462)	0.178*** (0.0483)	0.216*** (0.0474)
Distance		-0.846*** (0.0617)	-0.767*** (0.0648)	-0.361*** (0.137)	-0.900*** (0.0766)	-0.721*** (0.0675)	-0.784*** (0.0684)	-0.906*** (0.0539)	-0.771*** (0.0650)	-0.767*** (0.0647)
Contiguity		0.00882 (0.0779)	-0.197** (0.0848)	0.573*** (0.0880)	-0.126 (0.0812)	-0.277*** (0.0843)	-0.245*** (0.0848)	-0.256*** (0.0847)	-0.229*** (0.0864)	-0.197** (0.0849)
Country is an Island		0.770*** (0.0823)	0.523*** (0.0863)	0.291 (0.201)	0.477*** (0.0869)	0.415*** (0.0832)	0.494*** (0.0909)	0.463*** (0.0881)	0.504*** (0.0881)	0.526*** (0.0865)
lndlock_orig		-0.168 (0.109)	-0.0698 (0.122)	-0.0905 (0.195)	-0.0626 (0.121)	-0.0873 (0.122)	-0.111 (0.121)	-0.125 (0.121)	-0.0736 (0.123)	-0.0720 (0.123)
lndlock_dest		0.172* (0.0975)	-0.399*** (0.104)	-0.0418 (0.154)	-0.376*** (0.0984)	-0.475*** (0.110)	-0.418*** (0.105)	-0.419*** (0.106)	-0.452*** (0.110)	-0.399*** (0.104)

	Control Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
language_simasjp			1.288*** (0.147)	0.790*** (0.237)	-1.392** (0.555)	0.853*** (0.151)	1.253*** (0.155)	0.875*** (0.154)	1.238*** (0.145)	1.289*** (0.147)
Tourism flow				0.166** (0.0789)						
Distlanguage_simasjp					0.348*** (0.0715)					
Patent Orig						0.398*** (0.0458)				
Patent Dest						0.00715 (0.0585)				
SimilarPat							-0.0267 (0.0337)			
DistMultPt								0.179*** (0.0339)		
LogDifPtlanguage_simasjp									0.178*** (0.0656)	
crisis_2008										-0.122* (0.0688)
_cons	8.358*** (0.106)	1.582** (0.796)	2.610*** (0.820)	1.384 (1.512)	3.606*** (0.894)	0.318 (1.009)	3.054*** (0.819)	1.964** (0.903)	2.999*** (0.828)	2.641*** (0.817)
N	12618	12618	11543	1660	11543	10210	10210	10210	10210	11543
adj. R-sq										
Bic	329216989.7	101543450.7	75938267.8	4972635.4	75298564.6	68785192.4	71625000.9	69774118.4	71454729.4	75870851.7
Aic	329216892.9	101543287.0	75938098.6	4972505.5	75298388.1	68785011.7	71624827.4	69773944.9	71454555.8	75870675.2

Notes: Standard errors in parentheses; * p<0.10; ** p<0.05; *** p<0.01

Table 3: PPML Regression for FDI from High Income Countries

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Distance	-0.983*** (0.0710)	-0.848*** (0.0685)	-0.438*** (0.141)	-0.969*** (0.0838)	-0.866*** (0.0733)	-0.866*** (0.0687)	-0.990*** (0.0616)	-0.880*** (0.0709)	-0.848*** (0.0685)
language_simasjp		1.013*** (0.139)	0.768*** (0.252)	-1.354** (0.563)	0.566*** (0.153)	0.943*** (0.143)	0.650*** (0.159)	0.881*** (0.134)	1.011*** (0.139)
Tourism flow			0.228*** (0.0719)						
Distlanguage_simasjp				0.311*** (0.0737)					
Patent Orig					0.418*** (0.0532)				
Patent Dest					-0.0347 (0.0391)				
SimilarPat						0.00499 (0.0261)			
DistMultPt							0.152*** (0.0293)		
LogDifPtlanguage_simasjp								0.231*** (0.0671)	
crisis_2008									-0.118* (0.0681)
_cons	5.252*** (0.892)	5.302*** (0.890)	0.739 (1.764)	6.150*** (0.972)	3.114*** (1.145)	6.050*** (0.894)	4.554*** (1.017)	6.089*** (0.910)	5.337*** (0.887)
N adj. R-sq	10297	9337	1495	9337	8207	8207	8207	8207	9337
Bic	83604322.8	61033406.6	4156698.2	60576900.9	55728950.0	57676985.6	56738745.9	57343979.8	60971988.8
Aic	83604163.5	61033242.3	4156570.8	60576729.5	55728774.7	57676817.3	56738577.6	57343811.4	60971817.4

Notes: Table 3 has the same control variables as Table 2, so this table only presents the reduced version; standard errors in parentheses: * p<0.10; ** p<0.05; *** p<0.01

Table 4: PPML Regression for FDI from Low Income Countries

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Distance	-0.771*** (0.134)	-0.921*** (0.138)	0.465 (0.432)	-0.988*** (0.143)	-0.570*** (0.129)	-0.891*** (0.129)	-0.871*** (0.142)	-0.866*** (0.128)	-0.940*** (0.139)
language_simasjp		0.147 (0.611)	4.258** (2.029)	-3.795 (3.163)	0.356 (0.656)	-0.382 (0.666)	-0.0827 (0.677)	0.101 (0.642)	0.0717 (0.616)
Tourism flow			0.448** (0.218)						
Distlanguage_simasjp				0.476 (0.407)					
Patent Orig					0.407*** (0.110)				
Patent Dest					-0.402*** (0.0885)				
SimilarPat						-0.368*** (0.0358)			
DistMultPt							-0.0457 (0.0658)		
LogDifPtlanguage_simasjp								0.142 (0.356)	
crisis_2008									-0.806*** (0.257)
_cons	3.958** (1.801)	5.882*** (1.887)	-6.133 (8.786)	6.404*** (1.915)	3.416 (2.130)	4.938** (1.997)	6.004*** (2.105)	5.564*** (1.994)	6.468*** (1.894)
N adj. R-sq	2321	2206	165	2206	2003	2003	2003	2003	2206
bic Aic	7268938.8 7268812.3	6584019.5 6583888.4	143761.6 143693.3	6573214.6 6573077.8	5944591.3 5944451.2	5271576.2 5271441.7	6326306.2 6326171.8	6329871.7 6329737.2	6532548.6 6532411.8

Notes: Table 4 has the same control variables as Table 2, so this table only presents the reduced version; Standard errors in parentheses: * p<0.10; ** p<0.05; *** p<0.01