Determinants of outward FDI from emerging economies

JEL Classification: F23

Keywords: foreign direct investment (FDI), knowledge-capital (KC) model, emerging multinationals, Pseudo-Poisson maximum likelihood (PPML)

Abstract

Research background: The last four decades have witnessed an upsurge of multi-nationals from emerging markets alongside a narrowed gap in growth prospects between developed and emerging economies. UNCTAD statistics show that FDI flows from emerging economies have gone steady since 1980 and occupied more than one fifth of global FDI stock in 2015. Japan led the reverse FDI trend when it started to invest abroad in the 1960s and 1970s. Two decades later, in the 1980s-1990s, the reverse FDI trend was continued by so-called Asian tigers, then recently by those rapidly-industrializing economies in Southeast Asia as well as China and India in East and South Asia.

Purpose of the article: The main goal of this paper is to contribute empirically to the study of the determinants of FDI outflows from emerging economies.

Methods: In order to derive empirically testable hypotheses this paper refers to theoretical Knowledge-Capital model developed by Markusen (2002). The model is estimated using the Poisson-Pseudo Maximum Likelihood estimation technique. The specific research hypotheses derived from the theory are verified using a panel dataset of 38 home emerging countries and 134 host countries over the period 2001–2012.

Findings & Value added: In this paper, we distinguish between horizontal and vertical reasons for FDI. Our estimation results support the hypothesis that main-stream theory of multinational enterprise can explain FDI flows from emerging economies, implying the significant roles of total market size, skilled-labor abundance, investment cost, trade cost as well as geographical distance between two countries.
Introduction

The last four decades have witnessed an upsurge of multinational enterprises (MNEs) from emerging markets alongside a narrowed gap in growth prospects between the advanced and the less advanced economies. The UNCTAD statistics show that FDI flows from emerging economies have gone steady since 1980 and occupied more than one fifth of global FDI stock in 2015 (UNCTAD, 2017). Japan led the reverse FDI trend when it started to invest abroad in the 1960s and 1970s. Two decades later, in the 1980s–1990s, the reverse FDI trend was continued by so-called Asian tigers, then recently by those rapidly-industrializing economies in Southeast Asia as well as China and India in East and South Asia, respectively.

The reverse FDI trend also spreads to other continents and attracts a growing body of research into the drivers of this phenomenon. A number of theories and empirical studies have attempted to identify the determinants of reverse FDI. Yet the literature addressing this topic is still at an early stage of development and consists of a handful of studies on a very limited number of countries, hence not revealing the big picture. More importantly, most empirical studies were not based on any specific theoretical framework, so estimation results were hard to interpret. Therefore, the main hypothesis to be addressed in this paper is whether modern mainstream theories that explain FDI activity of MNEs from developed countries are also able to account for investment decisions of their counterparts from emerging economies or not.

The main goal of this paper is to contribute empirically to the study of the determinants of outward FDI from all the emerging economies. In order to derive empirically testable hypotheses, this paper employs the theoretical Knowledge-Capital (KC) model developed by Markusen (2002) which is subsequently estimated using the Poisson-Pseudo Maximum Likelihood (PPML) estimation technique. The specific research hypotheses derived from the KC model are validated using a panel dataset of 38 home emerging economies and 134 host countries over the period 2001–2012.

Our estimation results support the hypothesis that modern mainstream theories are able to explain FDI flows from the emerging economies, implying the significant roles of the total market size, skilled-labor abundance, investment cost, trade cost as well as geographical distance between home and host countries.

The structure of this paper is organized as follows. In the next section, we review the relevant FDI literature, focusing on investments made by MNEs originating from the emerging economies. Then, we explain the analytical framework of the Knowledge-Capital model and state the re-
search hypotheses derived from this model. Subsequently, we describe statistical data and the research methodology. Finally, we report and discuss our estimation results. The paper ends with concluding remarks and guidelines for further research.

**Literature review**

Emerging economies are often viewed in a relationship with other economic groups of countries. Compared to the developed markets, the emerging economies have less accessibility, but demonstrate a higher level of openness compared to the developing economies.\(^1\) In response to the growing importance of MNEs originating from the emerging markets in the world economy, substantial efforts have been made to explain their investment decisions in the economic and business literatures. Empirical studies look at this phenomenon from three broad perspectives, namely: the FDI perspective, the institutional perspective and the managerial perspective.

Unbound by the mainstream theories, the FDI perspective investigates an in-depth relationship between home and host country characteristics and the activity of MNEs. According to this perspective, outward and inward FDI positions of a country are strongly related to its level of economic development. This argument was proposed in the Dunning’s Investment Development Path (IDP) theory. Accordingly, to evolve from a net FDI recipient to a net investor, a country would need to go through five stages (Dunning, 1981, 1986; Dunning & Narula, 1996).

- **Stage 1:** At this stage, little inward and outward FDI takes place because country-level advantages are too few to attract FDI, with possible exceptions related to the extraction of natural resources. Local firms are not established yet or have not sufficiently developed their advantages to compete at an international level.
- **Stage 2:** At this stage, inward FDI starts to rise when per capita income increases and the local economy builds up its location advantages; while the level of outward FDI remains low and negligible.
- **Stage 3:** At this stage, the growth rate of outward FDI starts to increase while that of inward FDI is supposed to fall because local firms are competent enough to compete with foreign firms.

\(^1\) In fact, there is no uniform rule to classify economies in the world into the advanced economies, the emerging markets and the developing economies. Every international organization such as IMF or World Bank, for their research purposes, classifies economies in their own way. Fortunately, most databases present similar country groups, hence there would be hardly any problem when referring to studies using different databases.
Stage 4: At this stage, an important economic breakthrough occurs when outward FDI exceeds or equals inward FDI. Most domestic firms are experienced enough to compete with foreign firms in both home market and foreign markets.

Stage 5: At this stage, the stocks of inward and outward FDI are roughly equal, so net investment fluctuates around zero.

Moreover, home countries enable MNEs from emerging economies to compete in foreign markets with low prices (Pangarkar & Lim, 2003; Enderwick, 2009). In fact, their low-cost advantage may come from an easy access to natural resources (BCG, 2009) or other factor endowment resources such as cheap labor. This means that home country characteristics lay foundations for their firms’ expansion strategies. Hence, modern mainstream economic theories have taken into consideration only macroeconomic factors as the main determinants of MNEs’ investment decisions.

However, after some studies documented various market imperfections in the emerging countries, economists’ perception has shifted to other elements that may affect FDI flows, including institutional factors (Amal et al., 2009). Realizing a significant difference in the institutional context between the developed and the emerging countries, the international business literature has proposed an institutional perspective as an alternative perspective on the phenomenon of reverse FDI.

On the one hand, good institutions are expected to improve markets’ structure efficiency and motivate FDI activity (Mudambi & Navarra, 2002). Describing institutions as structures responsible for social behavior interaction between politics (such as corruption, transparency, etc.), law (such as economic freedom and regulatory regime) and society (such as ethical rules and business environment), Peng et al. (2008) sees the essential role of institutions in improving firms’ competitive advantages. McMillan (2007) argues that institutions are in fact present throughout strategy implementation and competitive development of local firms.

On the other hand, there is a possibility that negative institutional contexts may have a positive impact on FDI since emerging firms invest abroad to avoid unfavorable domestic investment climate (Cuervo-Cazurra & Genc, 2008). Unfavorable institutional contexts are interpreted as macroeconomic volatility, political instability, policy uncertainty, protectionism (Stal & Cuervo-Cazurra, 2011) and poor institutions such as weak property rights (Wu & Chen, 2014). The negative experience of emerging MNEs when operating in home markets has prevented them from developing their competitive advantages and driven them away. Such firm’s behavior is called by Luo et al. (2010) as institutional escapism.
The managerial perspective addresses differences and similarities in internationalization processes of MNEs from countries with different levels of economic development. Using behavioral approaches, this strand in the literature has pointed out specific features of internationalization patterns between the developed and the emerging economies:

- Emerging MNEs are based in countries with lower average per capita income and weaker institutional infrastructure;
- Emerging MNEs have limited ownership advantages, such as technology and brand when operating internationally;
- As latecomers, emerging MNEs follow different paths regarding the location choice. They invest not only in other emerging countries, but also in the developed countries (Ramamurti & Singh, 2009).

Hence, to overcome both the liability of foreignness and latecomers’ disadvantages, MNEs based in the emerging economies may select an audacious international strategy to quickly familiarize them with customers. Investors may acquire strategic assets and already established brands (Luo & Tung, 2007; Bonaglia et al., 2007). In such a case, FDI serves as a foreign market entry mode and the most valued strategic assets are R&D and networking. This strategy is also adopted by firms to avoid discrimination of consumers and governments in the developed markets, which may come from the assumptions that: (1) compared with products made in the developed countries, products produced by firms in the emerging economies are of inferior quality due to lower technological levels and weaker safety standards; (2) workers in emerging country firms suffer from ill-treatment, such as low wages, unsafe working conditions and inadequate labor rights and (3) MNEs from the emerging markets incur a higher cost of capital than MNEs from the developed markets because they are riskier, given poorer governance and macroeconomic instability (Cuervo-Cazurra & Ramamurti, 2015). Known as discrimination escape, this type of behavior and institutional escape combine into the escapist FDI theory (Stoian & Mohr, 2016).

Without referring to a specific theoretical framework, most empirical studies on MNEs from the emerging economies focus on providing evidence for all three aforementioned perspectives. One strand in the empirical literature aims at testing the role of home market in FDI activity of emerging MNEs. The studies investigate how macroeconomic and institutional factors in home countries influence firms’ decision to invest abroad. Factors such as GDP, exchange rate, trade and inflation are often included in empirical models to capture the effects of market size, trade openness and macroeconomic stability. Estimation results show unclear conclusions.
about the effect of GDP on FDI outflows (Frenkel et al., 2004; Kyrkilis & Pantelidis, 2005).

In response, researchers proposed the use of GDP per capita variable as a better indicator of economic development in the sense that richer customers have a preference for advanced products (Kyrkilis & Pantelidis, 2005; Faria & Mauro, 2009). Also, the relationship of outward FDI with both trade and exchange rate is unclear. The FDI-trade relationship is substitutional in the case of market-seeking projects, while complementary for cases of efficiency-seeking or resource-seeking projects (Swenson, 2004; Seo & Suh, 2006). Exchange-rate effects also depend on FDI’s nature. Market-seeking MNEs will be motivated by a high exchange rate (depreciated currency) while performance-seeking MNEs will be for lower production costs due to low exchange rates (Xing & Wan, 2006).

More recent studies made a further step by incorporating both traditional and non-traditional variables in empirical models. For example, Amal et al. (2009) found significant positive effects of education and globalization but the negative effect of economic freedom. The effect of economic freedom is controversial in the sense that it may act indirectly towards outward FDI by promoting inward FDI and improving firms’ competitiveness (Chittoor et al., 2008). Faria and Mauro (2009) added variables measuring financial development, human capital and governance. The promotional role of governance was supported by empirical evidence in China (Rasiah et al., 2010) and Brazil (Arbix, 2010).

The overview of studies on MNEs from the emerging economies suggests that empirical research either investigates separately different aspects of FDI or analyzes FDI activity in a specific country case or a country group case. Furthermore, most empirical studies do not have clearly specified theoretical underpinnings, which makes their estimation results hard to interpret. Therefore, in this paper we derive our estimating equations directly from the formal Knowledge-Capital model of multinational enterprise proposed by Markusen (2002) which to the best of our knowledge has not been employed so far in the context of the emerging economies.

**Research methodology**

The mainstream economic literature identifies two main reasons for FDI: market seeking and efficiency seeking (Markusen, 2013). According to the first one, FDI allows to overcome distance and lower costs of foreign markets access. Foreign direct investment undertaken with the aim to serve local markets is often called horizontal FDI. It refers to producing abroad

roughly the same goods and services as in the home country. According to
the second one, FDI is made in order to acquire inputs at a lower cost. For-
eign direct investment aiming at production cost reductions is called verti-
cal FDI. It involves international fragmentation of the value chain and lo-
cating various stages of production in different countries where the factors
used intensively in particular stages are relatively cheap.

The new trade theory (NTT) that emerged in the early 1980s provided
a set of modeling tools that proved useful in studying the determinants of
FDI. On the one hand, in order to explain FDI between countries at the
similar level of economic development a number of models of horizontally
integrated MNEs were developed. The early examples of this approach
include models developed by Krugman (1983) and Markusen (1984) that
were later extended, *inter alia*, by Horstmann and Markusen (1987),
Collie (2011) and Cieślik (2013; 2015a,b; 2016; 2018). On the other hand,
in order to explain FDI between the developed and developing countries
a number of models of vertically integrated MNEs were proposed. The first
models of a vertically-integrated multinational enterprise were developed
by Helpman (1984) and Helpman and Krugman (1985). These models were
later extended by, *inter alia*, Zhang and Markusen (1999), Markusen and

Initially, horizontal and vertical models of MNEs were treated as two
separate strands in the literature. The key step in the development of the
modern theory of the multinational enterprise was aimed at combining the
horizontal and vertical approaches into a hybrid framework in which firms
can choose between national, horizontal and vertical strategies. This was
done by Markusen (2002), who called this integrated framework the
knowledge-capital model. His model is currently regarded as the most gen-
eral theory of the multinational enterprise that allows national firms, hori-
zontal multinationals and vertical multinationals to emerge endogenously in
the equilibrium, depending on various combinations of home and host
country characteristics.

The KC model cannot be solved analytically and most results have to be
derived from numerical simulations.\(^2\) These simulations generate predic-
tions on the relationship between the extent of multinational activity and
home and host country characteristics. For example, national firms export-
ing to each other’s markets are the dominant type when countries are simi-
lar in economic size and relative factor endowments and trade costs are

\(^2\) The simulation results of KC model were demonstrated with a series of world Edge-
worth box diagrams.
low. In contrast, horizontal multinationals dominate when countries are similar in economic size and relative factor endowments but trade costs are high. However, if countries are dissimilar in either size or in relative factor endowments one country is favored as a location of both headquarters and production activities or one of these two activities.

In particular, if countries are dissimilar in size but similar in relative factor endowments, then national firms located in the large country are favored, as they can avoid installing costly capacity in the smaller market. On the other hand, if countries are similar in size but dissimilar in relative factor endowments, vertical multinationals are the dominant type, as there is an incentive to split the production process and locate headquarters in the human-capital abundant home country and production in the labor-abundant host country, unless trade costs are high. The extent of multinational activity in the KC model is the largest when the home country is moderately small and highly abundant in human capital.

Although most findings of the KC model are derived from numerical simulations, the model generates a number of testable predictions, related to the extent of multinational activity to country characteristics. The bilateral relationships between firm types and economic characteristics of two countries: country $i$ and country $j$, derived from the KC model are illustrated in Table 1.

The predictions of the KC model can be tested using statistical data on FDI from the emerging countries. We assume that MNEs are headquartered in the emerging home countries, which means that they are the i-country and all other countries (including the emerging ones) are the j-country in the theoretical model. Our research hypotheses on outward FDI from the emerging markets derived on the basis of the KC model can be formulated as follows:

Hypothesis 1: *Total income and the similarity in market size between home and host countries are associated positively with FDI (horizontal reason).*

Hypothesis 2: *The differences in relative factor endowments between home and host countries are positively associated with FDI (vertical reason).*

Hypothesis 3: *Higher investment freedom in host country leads to higher FDI (both horizontal and vertical reasons).*

Hypothesis 4: *High trade costs between home and host countries discourage vertical FDI but encourage horizontal FDI so the overall effect is not clear.*
In addition, we also include some other typical factors usually shown having effects on FDI in previous empirical studies, namely the common language between host and home countries. This variable may affect transaction costs in doing business (Kim et al., 2014). Thus, we expect those variables to be also positively related to FDI.

Hypothesis 5: Common spoken language encourages FDI (both horizontal and vertical reasons).

As it has not been clear from the literature which measure of FDI is best this paper uses both FDI outflows and stocks from emerging markets as the dependent variable. UNCTAD provides annual data on bilateral FDI flows and stocks for almost every country for period 2001–2012. Thus, for each year one country may serves as both a home country for its firms to produce abroad and a host country for foreign firms to produce there.

With 38 emerging home countries and 134 host countries for 12-year period, the highest possible number of observations in our sample would be 61,104. However, due to data unavailability of outward FDI flows and stocks for some country pairs in some years, the sample size ends up with a much lower number of observations. Annual FDI flows and stocks are converted into millions of 2011 US dollars using host country’s GDP deflator taken from World Bank data.

In the empirical implementation of the KC model, both FDI outflows and stocks are estimated using the following equation:

\[
\text{OFDI}_{ijt} = \beta_0 + \beta_1 \text{lnsumGDP}_{ijt} + \beta_2 \text{lnsimilarity}_{ijt} + \beta_3 \text{SKd}_{ijt} + \\
\beta_4 \text{ldist}_{ij} + \beta_5 \text{Language}_{ij} + \beta_6 \text{lnINVC}_{ijt} + \beta_7 \text{ITC}_{it} + \beta_8 \text{ITC}_{jt} + \nu_t + \varepsilon_{ijt}
\]

where: \(\text{OFDI}_{ijt}\) is the measure of foreign direct investment from home country \(i\) in host country \(j\) in year \(t\), \(\text{lnsumGDP}_{ijt}\) is the natural log of the sum of GDPs of country \(i\) and country \(j\) in year \(t\), \(\text{lnsimilarity}_{ijt}\) is the natural log of the measure of similarity in market size between countries \(i\) and \(j\) in year \(t\), \(\text{SKd}_{ijt}\) is the difference in skilled labor intensity between country \(i\) and country \(j\) in year \(t\), \(\text{ldist}_{ij}\) is the natural log of geographic distance between the most populated cities in countries \(i\) and \(j\), \(\text{Language}_{ij}\) is the binary variable taking value 1 if countries \(i\) and \(j\) share the official language and 0 otherwise, \(\text{lnINVC}_{ijt}\) is the natural log of the investment cost in host country \(j\) in year \(t\), \(\text{ITC}_{it}\) is the natural log of the trade cost for goods exported to home

---

3 Table 2 lists the names of host countries and their economic status.
4 Detailed explanations, summary statistics and correlations between explanatory variables are provided in tables 3, 4 and 5.

---

217
country i in year t, ITC_{jt} is the natural log of the trade cost for goods exported to host country j in year t, \(v_t\) is the individual time effect for each year, and \(\varepsilon_{ijt}\) is the error term.

Our dependent variables OFDI i.e. bilateral FDI flows and stocks between two countries, assume non-negative values. This characteristic challenges the traditional estimation technique, such as OLS for example, in the sense that the straightforward linear OLS is incapable of guaranteeing nonnegative predicted values of the dependent variable. There usually will be values of explanatory variables such that the predicted value of OFDI is negative.

The negativity problem can be solved by using natural log transformation, which in this case is \(\ln(\text{OFDI})\) then using a linear model. However, this approach is only applicable for a strictly positive dependent variable, whereas in our sample OFDI can be literally zero. For example, the flow of FDI between some country pairs was interrupted for some years after 2008’s global economic crisis. Several methods have been developed to address this problem, most commonly deleting zero observations from the dataset, then log-transforming the rest and estimating by OLS. To avoid losing information, some authors replace zero values with a small positive constant (an arbitrarily small value) then run the regression on the new dependent variable with Tobit estimator. Nevertheless, without any strong theoretical or empirical justification, both methods will more or less distort estimation results. Besides, the log-transformation itself is heavily criticized for causing Jensen’s inequality and inconsistent coefficient estimates in the presence of heteroskedasticity (Silva & Tenreyro, 2006).

Over the years, efforts have been taken to find an appropriate alternative estimation technique to deal with problems of zero dependent variables and log-transformation. As a consequence, a large number of estimators are employed, i.e. Poisson and modified Poisson models, Nonlinear Least Squares, Feasible Generalized Least Squares (FGLS) and Helpman et al. (2008) approach. Among them, Poisson-Pseudo Maximum Likelihood (PPML), which is proposed by Silva & Tenreyro (2006) stands out as best performing estimator. Though later found out by Martinez-Zarzaso (2013) that it is not always the best estimator, and sometimes outperformed by both OLS and FGLS in the out-of-sample forecast, PPML, with its identified advantages can be considered as a benchmark against which other alternative estimators can be compared (Silva & Tenreyro, 2013). So, this study determines to estimate KC model with the PPML estimator.

PPML does not take logs and estimates the model in levels, hence avoiding log-transformation problems. To be specific, PPML estimates \(\beta\) in
Specification (1) by maximizing the probability of observing a count of OFDI$_{ijt}$, with the function as below:

$$
Pr(\text{OFDI}_{ijt}|X) = \frac{e^{-\lambda_{ijt}X} \lambda_{ijt}^{\text{OFDI}_{ijt}}}{\text{OFDI}_{ijt}!} 
$$

where: $\lambda_{ijt}$ is the expected value of OFDI, assumed to be log-linearly dependent on the vector of explanatory variables $X_{ijt}$: $\ln \lambda_{ijt} = X_{ijt}\beta$

By log-transformation, we obtain the log-likelihood function:

$$
L(\beta) = -\sum \ln \lambda_{ijt} + \sum \text{OFDI}_{ijt} \ln(\lambda_{ijt}) - \sum \ln(\text{OFDI}_{ijt}!)
$$

$$
= -\sum \exp(X_{ijt}\beta) + \sum \text{OFDI}_{ijt} (X_{ijt}\beta) - \text{constant}
$$

The maximization of $L(\beta)$ requires solving the following first- and second-order conditions:

$$
\frac{dL}{d\beta} = -\sum [\exp(X_{ijt}\beta) - \text{OFDI}_{ijt}] X_{ijt} = 0 
$$

and

$$
\frac{d^2L}{d\beta d\beta'} = -\sum [\exp(X_{ijt}\beta)] X_{ijt} X' < 0
$$

As long as the conditional mean is correctly specified: $E[\text{OFDI}_{ijt}|X] = \exp(X_{ijt}\beta)$, $\beta$ will be consistently estimated. In addition to this robustness property, PPML estimator is well-behaved in the sense that the second-order condition is easily satisfied for all $X$ and $\beta$, hence facilitating the estimation and ensuring the uniqueness of the maximum.

One note to the use of PPML estimator is its equidispersion assumption. Accordingly, the conditional mean $E[\text{OFDI}_{ijt}|X]$ given as $\exp(X_{ijt}\beta)$ is equal to the conditional variance $V[Y_{ijt}|X]$. So restrictions are imposed on conditional moments of the OFDI as below:

$$
E[\text{OFDI}_{ijt}|X] = \exp(X_{ijt}\beta) \propto V[\text{OFDI}_{ijt}|X]
$$

However, this assumption is often unlikely to hold as the estimator does not fully account for the presence of heteroskedasticity in the model. Specifically, the estimator does not cover unobserved heterogeneity and makes
the conditional variance greater than the conditional mean (also known as overdispersion). Hence, all reference has to be based on an Eicker-White robust covariance matrix estimator (Eicker, 1967; White, 1980). In response, Silva and Tenryro (2011) wrote a Stata command starting with `ppml` to execute the regression, so the implementation of robust PPML estimator is quite straightforward. Moreover, different from Stata’s built-in `poisson` command, `ppml` identifies and drops regressors that cause the nonexistence of the pseudo-maximum likelihood estimates in Poisson regression, hence guaranteeing that model assumptions are satisfied and relevant variables are selected.

In short, with the following properties, PPML claims to be a promising workhorse in estimating the equation as:

- PPML can produce unbiased and consistent estimates, robust to different patterns of heteroskedasticity. It avoids under-prediction of large FDI volume by generating estimates of FDI flows rather than the log of FDI flows.
- All that is needed for PPML estimator to be consistent is the correct specification of the conditional mean: $E[\text{OFDI}_{ijt}|X] = \exp(X_{ijt}\beta)$. Data do not have to follow Poisson distribution and more importantly, dependent variable OFDI does not have to be an integer (Gourieroux et al., 1984)
- PPML method well behaves even in the presence of overdispersion in the dependent variable and large proportion of zeros in the sample (Silva & Tenryro, 2011).

Therefore, PPML is the optimal estimator for the equation and will be employed in our study.

**Results**

In this section we report two sets of our empirical results. First, in column (1) of Table 6 we report estimation results for KC model where the dependent variable is FDI outflow from emerging investor countries. Then in column (2) of Table 6 we report estimation results for KC model where the dependent variable is FDI stock. As stated in our main research hypothesis, regression results are expected to follow the KC model’s predictions. For comparison, there is an additional column showing expected signs on the estimated coefficients. Table 6 presents our estimation results of the KC model. For brevity, time-dummy variables are not reported (available on request).

As can be seen from column (1) in Table 6, almost all estimated coefficients are statistically significant at the 1% significance level and display
the expected signs. The statistically significant estimated coefficients on $l\text{sumGDP}$, $l\text{similarity}$, $SK_d$ and $l\text{TC}_i$ confirm our research hypotheses and provide evidence for both horizontal and vertical FDI. In particular, total income and similarity in market size between emerging investor and recipient country motivate horizontal FDI, while differences in skilled-labor abundance motivate vertical FDI. In fact, this is among few studies that demonstrate the co-existence of both FDI types.

With regards to other explanatory variables in the model, it is convenient to classify them into two groups related to, either FDI enhancement (i.e. Language) or FDI friction (i.e. $l\text{dist}$, and $l\text{INVC}_j$). In short, the KC model explains relatively well outward FDI flows from the emerging economies and provides evidence for both FDI types.

Moreover, the estimated coefficient on $l\text{TC}_j$ is negative and statistically significant at the 1% level, implying that costs of exporting goods to host country is a hindrance to FDI activity. These results go against horizontal FDI motivation that higher trade costs would discourage trade and encourage FDI activity and firms would choose to enter the foreign market via FDI rather than exports. A probable explanation of this result is that in reality the vertical reason for FDI might be stronger than the horizontal one. Moreover, horizontally-integrated MNEs may still need to transport some intermediate inputs from home country and would also prefer low trade costs.

Column (2) reports the estimation results obtained from the alternative specification of the KC model in which the dependent variable is the outward FDI stock instead of FDI outflow. These estimation results are very similar in qualitative terms to the results reported in column (1). All the explanatory variables are statistically significant and display the expected signs. Hence, irrespectively of the used measure of FDI the empirical results support the predictions of the KC model.

**Conclusions**

This empirical study explored the determinants of outward FDI flows from the emerging economies that differ in terms of economic and institutional contexts from the developed countries. In particular, we investigated whether the mainstream economic theory that successfully explains FDI activity of MNEs from the developed economies is also able to account for investment decisions of their emerging market counterparts. To validate the explanatory power of the mainstream theory of multinational enterprise, we estimated Markusen’s Knowledge Capital model that combines horizontal
and vertical reasons for FDI using the PPML technique on a panel data set of FDI flows from 38 emerging countries to 134 host countries for the period 2001–2012. In line with this model, market size, skilled-labor abundance, investment costs, trade costs, as well as bilateral geographic distance, were found to be significant determinants of outward FDI from the emerging economies.

Our study opens a fruitful avenue for future research on FDI activity of emerging market investors. Future research may go in one of several directions. First, instead of lumping all sectors in one framework, follow-up studies may disaggregate data by sector and investigate determinants of FDI using sector-level data. Some studies have already taken this approach and argued that MNEs’ choice of FDI type may vary across sectors. Though requiring a lot of data input and heavily dependent on data availability, this approach is rewarding in the sense that it avoids aggregation bias and generates sector-specific insights. It may contribute to the literature that mostly focuses on aggregated sectoral classifications or on the manufacturing sector only. Second, empirical studies can take a further step by separately investigating FDI activity of emerging investors in host countries of different levels of economic development. Based on the same empirical framework of the KC model, future studies can determine which FDI type (or both) dominates in different destinations, hence offering interesting insights into investment behavior of MNEs from the emerging markets.

Finally, also depending upon data availability, future studies can broaden time period from 2001 backwards or 2012 forwards to capture a longer period and use the time-series dimension of FDI flows. The current study used dummy variables as proxies for the presence of a third country in the existing two-country framework. From the technical point of view, future studies can develop new econometric tools or improve existing estimation techniques to handle the likely appearance of negative FDI flows (divestment) in future datasets. Though the phenomenon is only emerging under certain circumstances and not significant enough to be noticed, an improved estimation technique would fully utilize the available data and generate more economic insights, for example reflect unfavorable investment climate during certain periods of time.
References


Annex

Table 1. Firm types and countries’ characteristics in KC model

<table>
<thead>
<tr>
<th>Dominant firm type</th>
<th>Different in size and relative endowment</th>
<th>Similar in size, relative endowment, factor prices</th>
<th>Trade cost</th>
<th>Total income</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Horizontal firms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type-hi</td>
<td>Not high foreign investment barriers</td>
<td>Yes</td>
<td>High</td>
<td>High</td>
<td>Type-hj will also produce in i</td>
</tr>
<tr>
<td>Type-hj</td>
<td>Not high foreign investment barriers</td>
<td>Yes</td>
<td>High</td>
<td>High</td>
<td>Type-hi will also produce in j</td>
</tr>
<tr>
<td><strong>Vertical firms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type-vi</td>
<td>- Small</td>
<td>Not high foreign investment barriers</td>
<td>Not excessive</td>
<td>Low</td>
<td>Trade costs here are costs from the host country back to the home country</td>
</tr>
<tr>
<td>Type-vj</td>
<td>Not high foreign investment barriers</td>
<td>Not high foreign investment barriers</td>
<td>Not excessive</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td><strong>National or domestic firms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type-di</td>
<td>- Skilled-labor abundant</td>
<td>High foreign investment barriers</td>
<td>Yes</td>
<td>Low</td>
<td>Type-di may sell in j if trade cost is not excessive</td>
</tr>
<tr>
<td>Type-dj</td>
<td>High foreign investment barriers</td>
<td></td>
<td>Yes</td>
<td>Low</td>
<td>Type-dj may sell in i if trade cost is not excessive</td>
</tr>
</tbody>
</table>

Source: own summary based on Markusen (2002).
Table 2. List of host countries classified by their economic status

| Developed economies | Australia, Austria, Canada, Cyprus, Demark, Estonia, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Malta, Netherlands, New Zealand, Norway, Portugal, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States |
| Emerging economies | Argentina, Bahrain, Brazil, Chile, China, Colombia, Croatia, Czech Republic, Ecuador, Egypt, Hong Kong, Hungary, India, Indonesia, Kazakhstan, South Korea, Kuwait, Malaysia, Mexico, Nigeria, Oman, Peru, Philippines, Poland, Qatar, Romania, Russia, Saudi Arabia, Singapore, Slovakia, South Africa, Thailand, Turkey, Ukraine, United Arab Emirates, Uruguay, Venezuela |
| Developing & least developed economies | Albania, Algeria, Armenia, Azerbaijan, Bahamas, Bangladesh, Barbados, Belarus, Belize, Benin, Bolivia, Bosnia and Herzegovina, Bulgaria, Cambodia, Cameroon, Congo, Costa Rica, Côte d’Ivoire, Dominican Republic, El Salvador, Fiji, Gabon, Georgia, Ghana, Guinea, Haiti, Honduras, Iran, Iraq, Jamaica, Jordan, Kenya, Kyrgyzstan, Lao, Lebanon, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Moldova, Morocco, Mozambique, Nepal, Nicaragua, Niger, Pakistan, Panama, Paraguay, Rwanda, Senegal, Sierra Leone, Sri Lanka, Sudan, Suriname, Syrian Arab Republic, Tajikistan, The FYR of Macedonia, Togo, Tunisia, Uganda, United Rep. of Tanzania, Uzbekistan, Viet Nam, Yemen, Zambia, Zimbabwe |

Source: own summary from UNCTAD (2016).

Table 3. Description of variables in KC model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Expected sign</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFDI</td>
<td>FDI outflows/stocks from home country i to host country j in a certain year (in million 2011 USD)</td>
<td>UNCTAD</td>
<td></td>
</tr>
<tr>
<td>lsumGDP</td>
<td>Log of sum of real GDP of both countries (in million 2011 USD)</td>
<td>+</td>
<td>Penn World Tables 9.0</td>
</tr>
<tr>
<td>lsimilarity</td>
<td>liksimilarity = log [(GDP_i / (GDP_i + GDP_j)) x (GDP_j / (GDP_i + GDP_j))]</td>
<td>(Horizontal motivation for FDI activity)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>When country i and j are identical in size (GDP_i = GDP_j = ½ of sumGDP), similarity is maximized (= ¼)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SKd</td>
<td>Difference in skilled labor intensity between country i and country j - Skilled labor intensity is defined as share of high skilled workers (level 3&amp;4 according to International Standard Classification of Occupations ISCO-08) in total labor force, thus having a potential range of 0 (very skilled-labor scarce) to 1 (very skilled-labor abundant) - SKd is positive if home country i is skilled-labor abundant relative to host country j</td>
<td>+</td>
<td>International Labor Organization</td>
</tr>
<tr>
<td>ldist</td>
<td>Log of geographical distance between home and host countries (between most populated cities, in km)</td>
<td>-</td>
<td>CEPII</td>
</tr>
</tbody>
</table>

Distance increases costs of investment
Table 3. Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Expected sign</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>Language is a binary variable (=1 if both countries share official language, =0 otherwise)</td>
<td>+</td>
<td>(Common language is likely to enhance information flows)</td>
</tr>
<tr>
<td>$\ln \text{INVC}_j$</td>
<td>- Log of investment cost in host country $j$ ($\text{INVC}_j$) [ \text{INVC}_j = 100 - \text{Investment Freedom Index} ] - Investment Freedom Index is computed on a scale from 0 to 100, with a larger number indicating more freedom. - The index evaluates a variety of restrictions that are typically imposed on investment. Points are deducted from the ideal score of 100 for each of the restrictions found in a country’s investment regime.</td>
<td>-</td>
<td>(Larger numbers indicate higher costs)</td>
</tr>
<tr>
<td>$\ln \text{TC}_i$</td>
<td>- Log of trade cost for goods exported to home country $i$ ($\text{TC}<em>i$) [ \text{TC}<em>i = 100 - \text{Trade Freedom Index} ] - Trade Freedom Index runs from 0 to 100, with 100 being the highest freedom level. - The index is a composite measure of the absence of tariff and non-tariff barriers that affect imports and exports of goods and services. - Trade Freedom index in country $i$ is computed as: [ \left( \frac{\text{Tariff}</em>{\text{max}} - \text{Tariff}</em>{\text{min}}}{\text{Tariff}<em>{\text{max}} - \text{Tariff}</em>{\text{min}}} \right) \times 100 - \text{NTB}<em>i ] Tariff$</em>{\text{max}}$ and Tariff$_{\text{min}}$ represent upper and lower bounds for tariff rates (%) Tariff, represents weighted average tariff rate (%) NTB, represents non-tariff barrier penalty</td>
<td>-</td>
<td>(Larger numbers raise costs of shipping goods back to home country from a foreign plant – vertical FDI motivation) The Heritage Foundation</td>
</tr>
<tr>
<td>$\ln \text{TC}_j$</td>
<td>Log of trade cost for goods exported to host country $j$ ($\text{TC}_j$) [ \text{TC}_j = 100 - \text{Trade Freedom Index} ]</td>
<td>+</td>
<td>(Higher trade costs encourage inward investment in host country)</td>
</tr>
<tr>
<td>$\nu_t$</td>
<td>Time dummy variable for each year from 2002-2012 (except 2001, which is the base year)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 4. Summary statistics of variables in KC model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFDI</td>
<td>138.12</td>
<td>1325.80</td>
<td>-13957.78</td>
<td>50501.03</td>
</tr>
<tr>
<td>lGDP</td>
<td>14.13</td>
<td>1.17</td>
<td>10.36</td>
<td>17.24</td>
</tr>
<tr>
<td>lsimilarity</td>
<td>-2.46</td>
<td>1.20</td>
<td>-8.73</td>
<td>-1.39</td>
</tr>
<tr>
<td>SKd</td>
<td>-0.03</td>
<td>0.16</td>
<td>-0.47</td>
<td>0.51</td>
</tr>
<tr>
<td>ldist</td>
<td>8.33</td>
<td>0.99</td>
<td>4.09</td>
<td>9.89</td>
</tr>
<tr>
<td>Language</td>
<td>0.15</td>
<td>0.35</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lINVc_j</td>
<td>3.57</td>
<td>0.57</td>
<td>1.61</td>
<td>4.60</td>
</tr>
<tr>
<td>lTC_i</td>
<td>3.23</td>
<td>0.44</td>
<td>1.61</td>
<td>4.36</td>
</tr>
<tr>
<td>lTC_j</td>
<td>3.13</td>
<td>0.46</td>
<td>1.61</td>
<td>4.61</td>
</tr>
</tbody>
</table>

Table 5 Correlations between variables of KC model

<table>
<thead>
<tr>
<th>Variable</th>
<th>OFDI</th>
<th>lsumGDP</th>
<th>lsimilarity</th>
<th>SKd</th>
<th>ldist</th>
<th>Language</th>
<th>lINVc_j</th>
<th>lTC_i</th>
<th>lTC_j</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFDI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lGDP</td>
<td>0.1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lsimilarity</td>
<td>-0.04</td>
<td>-0.42</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SKd</td>
<td>0.02</td>
<td>-0.3</td>
<td>-0.002</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ldist</td>
<td>-0.04</td>
<td>0.35</td>
<td>-0.18</td>
<td>-0.11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>0.11</td>
<td>-0.16</td>
<td>0.04</td>
<td>0.09</td>
<td>-0.13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lINVc_j</td>
<td>-0.02</td>
<td>0.01</td>
<td>-0.11</td>
<td>0.45</td>
<td>-0.07</td>
<td>0.05</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lTC_i</td>
<td>-0.05</td>
<td>0.25</td>
<td>-0.2</td>
<td>-0.3</td>
<td>0.19</td>
<td>0.07</td>
<td>0.05</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>lTC_j</td>
<td>-0.04</td>
<td>-0.08</td>
<td>-0.14</td>
<td>0.46</td>
<td>0.04</td>
<td>0.11</td>
<td>0.56</td>
<td>0.20</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: own calculations using STATA14.

Table 6. Estimation results of KC model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected sign</th>
<th>OFDI flow (1)</th>
<th>OFDI stock (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lsumGDP</td>
<td>+</td>
<td>1.371***</td>
<td>1.456***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.064)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>lsimilarity</td>
<td>+</td>
<td>0.147**</td>
<td>0.292***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.065)</td>
<td>(0.087)</td>
</tr>
<tr>
<td>SKd</td>
<td>+</td>
<td>2.989***</td>
<td>3.527***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.350)</td>
<td>(0.377)</td>
</tr>
<tr>
<td>ldist</td>
<td>-</td>
<td>-0.858***</td>
<td>-0.803***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.055)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Language</td>
<td>+</td>
<td>1.957***</td>
<td>2.110***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.102)</td>
<td>(0.117)</td>
</tr>
<tr>
<td>lINVc_j</td>
<td>-</td>
<td>-0.760***</td>
<td>-0.827***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.114)</td>
<td>(0.121)</td>
</tr>
</tbody>
</table>
### Table 6. Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected sign</th>
<th>OFDI flow (1)</th>
<th>OFDI stock (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITC₁</td>
<td>-</td>
<td>-0.508***</td>
<td>-1.007***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.121)</td>
<td>(0.155)</td>
</tr>
<tr>
<td>ITC₂</td>
<td>+</td>
<td>-0.392***</td>
<td>-0.417***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.140)</td>
<td>(0.140)</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>-2.896***</td>
<td>-2.896***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.664)</td>
<td>(0.146)</td>
</tr>
</tbody>
</table>

Number of observations: 14,177 (1), 14,057 (2)

Pseudo log-likelihood: -3,442,358.7 (1), -17,401,846 (2)

Pseudo R-squared: 0.520 (1), 0.569 (2)

*Note: Robust standard errors in parentheses; Significance level: (*** = 1%, (**) = 5%, (*) = 10%)*

Source: own computations using STATA14.