# Cambridge Working Papers in Economics

*Evolution of Bilateral Capital Flows to Developing Countries at Intensive and Extensive Margins*

*Juliana Araujo, Povilas Lastauskas and Chris Papageorgiou*

# Evolution of Bilateral Capital Flows to Developing Countries at Intensive and Extensive Margins[*]

Juliana Araujo[†]
International Monetary Fund

Povilas Lastauskas[‡]
University of Cambridge

Chris Papageorgiou[§]
International Monetary Fund

February 7, 2015

**Abstract**

The capital flows network has changed substantially, bringing new investors and target economies into play. Related, a recent intensification of capital flows to low income countries (LICs) has posed a number of questions. Most importantly, the very nature of those flows and important factors affecting foreign investors decision which can ultimately affect growth prospects of low income countries (together with an issue of sustainability) remain open for an academic probe. Due to an existence of a share of costs which is fixed in nature, there is a need to analyze capital flows and their evolution at two margins: intensive and extensive. This paper presents a parsimonious theoretical account that is consequently mapped into an econometric framework where we allow for two-tier decisions and cross-sectional dependence. Results indicate that market entry costs affect investment decisions pertinent to the LICs, consistently with the static theory. However, persistence in extensive margin eliminates this effect once dynamics is allowed for.

**Keywords**: Bilateral Capital Flows, Foreign Direct Investment, Portfolio Flows, Developing Economies, Extensive and Intensive Margins, Heterogeneous Panels, Cross-Sectional Dependence, Copulae.
**JEL Classification**: C33, C34, F21, F62, O16

## 1 Introduction

Despite capital being one of major manifestations of globalization and development engines, we lack understanding on what determines capital flows to low income countries. In this paper we consider two main types of private long-term capital movements, namely portfolio[1] and direct investment, though emphasize the latter.[2] The major difference between the two types is that direct investor has, on a lasting basis, management interest in a nonresident enterprise, whilst portfolio investment is purely financial in nature.[3] The stage of development of a country also influences the investors' decisions. All these issues

---

[†]Address: International Monetary Fund, Washington, DC 20431, USA. Email: jaraujo@imf.org.

[‡]Corresponding author. Address: Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge, CB3 9DD, UK. Email: pl312@cam.ac.uk.

[§]Address: International Monetary Fund, Washington, DC 20431, USA. Email: cpapageorgiou@imf.org.

[1] For example, holdings of government securities and private equities, with the latter significantly more important than the former, preference shares, equities without control over organization, bonds issued by international organizations, etc.

[2] Other types include international loans and commercial credits with a maturity of more than one year.

[3] Portfolio investment is traditionally considered to be a function of differential yields and risk diversification. Usually the set of efficient portfolios is first determined, independently of the stock of wealth, and later the optimum portfolio is determined employing the investor's utility function. Evidently, the yields, risks and tastes (and the stock of wealth itself), underlying the optimum composition, must be time-variant for the continuous flows of portfolio investment to take place.

require adequate theoretical and empirical accounts, paying attention to fixed costs, generating two-tier decisions, and flexible structure between them.

Moreover, the foreign direct investment deserves discussion with respect to real and financial aspects of investment. One way to analyze the FDI is to consider movements of the productive factor (real capital) between countries. The trade theory is then fully compatible: the causes of capital flows include different rewards in different economies due to prices, taxes and tariffs[4] and other reasons like specialization or factor-intensity reversals. Industry-specific and real economy factors shall be considered because usually big corporations with a high product differentiation engage in FDI.[5] The potential advantages include the transfer of entrepreneurship and new technology, whereas disadvantages assume outflows of repatriated profits, diminution in sovereignty, and economic policy inefficacies.[6] Razin and Sadka (2007a) carefully differentiate between the two, FDI and FPI, flows. FDI firms are more efficient at the expense of initial fixed cost and higher selling cost (lower liquidity), whereas others whose probability of a liquidity shock is higher choose FPI. Razin and Sadka (2007a) stress two margins of FDI decisions, driven by fixed setup costs. Maintaining wages fixed in the host country, a positive productivity shock in this country increases the marginal products of the factors of production (including capital). Hence, a positive effect on the flows of FDI that are governed by the intensive margin. However, when wages are allowed to adjust, the productivity shock generates an upward pressure on wages which raises the fixed setup costs and discourages FDI through the extensive margin.

We model FDI flows drawing from these insights in the unifying framework of a gravity model making use of bilateral (rather than aggregate) capital flows data. Our contribution revolves around three parts. First, a single parsimonious framework for both, FDI and portfolio, types of capital flows resulting in simple gravity-type relations is presented. Second, our framework enables us to analyze the changing structure of international investment along two, intensive and extensive, margins. Our simplifying assumption of separability allows telling supply story for FDI and demand story for portfolio investment, also admits a simple mapping from theory to empirics. Unlike, for example, Razin and Sadka (2007b,a) which develop a more dynamic approach for FDI by using current value functions for firms to determine whether an investment is worthwhile to take place. However, that approach results in a reduced form estimation, with less parsimony, weaker links between theory and empirics, and different emphasis (authors mainly rely on OECD countries with the longer time dimension). Third, by decomposing the sample into least developed countries and middle income economies, we are able to track differences across development stages. We offer new perspectives for direct and portfolio capital flows, and countries just starting to be actors in capital markets. Our parsimonious econometric framework allows for cross sectional dependencies, directional fixed and time effects, also can be cast in copula setting to avoid distributional assumptions. Before we turn to analyzing the changing nature of capital flows to LICs and the economic determinants of investors' decision, we first review the gravity framework and its applications to capital movements.

The paper is organized as follows. We overview theoretical and empirical contributions to FDI and portfolio literatures in Section 2. Section 3 depicts a number of recent stylized facts on developing countries, push and pull factors, intensive and extensive margins, and evolution of capital flows network. A theoretical model is laid down in Section 4. We discuss econometric setting in Section 5 and overview baseline results in Section 6. Extensions and robustness checks are undertaken in Section 7. Finally, Section 8 concludes and outlines future research directions and policy implications whereas Appendix collects all the supporting material.

---

[4] Stolper-Samuelson theorem can be invoked to determine a relationship between tariffs and factor rewards. As emphasized by Gandolfo (2002), the effects on the host country can be analyzed drawing from the Rybczynski's theorem because it addresses the effects of changes in factor endowment (capital stock).

[5] See Helpman, Melitz, and Yeaple (2004), for a formalization of heterogeneous multinationals engaging in production abroad as opposed to trade.

[6] Examples of a loss of sovereignty include instructions from the parent company which need to be followed instead of those from local authorities whereas monetary policy can become ineffective due to the resources to the financial market of the parent company.

## 2 Brief Literature Review

Though literature on capital flows is vast, contributions on the links between theory and empirics are scarce. We overview the use of gravity framework to model both FDI and portfolio flows. Empirical literature has focused on advanced economies and developing countries (see Levy Yeyati, Stein, and Daude (2003) among others) but with little if at all work trying to distinguish between developed countries at different stages of development. Though theory is pointing to capital markets efficiency and integration, also institutional quality to attract investment, we lack knowledge on factors which generate success or failure to be a target economy for international investors. Before dealing with specificities of low and middle income countries, we cover main theoretical and empirical determinants to engage in foreign investment as entertained by current literature.

### 2.1 Gravity in Direct Investment Flows

Foreign Direcet Investment (FDI) is usually modeled as an investment made by the multinational firms. The gravity-type dependence for the foreign affiliates' sales was derived by Kleinert and Toubal (2010). They employ three different models: the emergence of horizontal multinational firms with firm symmetry and heterogeneity, and the emergence of vertical multinational firms using a factor-proportions approach. Kleinert and Toubal (2010) assume that that fixed costs increase with distance, thereby leading to a negative relationship between distance and affiliate sales (the standard proximity - concentration model predicts an opposite result). The factor proportions model augments the otherwise standard gravity specification with a relative factor endowment and a joint size of home and host country regressor.

Obstfeld and Rogoff (2000) also argued that transportation costs may be relevant not only for commodity trade but also for international asset transactions. Mody (2007) collects research on what attracts and what discourages investment in a foreign country. Distances, especially informational and transactional do play a role as does "herding" behavior believed to be rooted in informational advantages of previous investors into the country. Another line of inquiry is concerned with benefits of FDI. It is found that FDI is not integrating the world in the sense that it flows to places where conditions are already propitious. Interestingly, Mody and Murshid (2005) find that developing countries with better policies succeeded more in absorbing foreign inflows by creating an environment conducive for the diffusion of new technologies and ideas intrinsic to foreign capital. Improved policies probably also reduced the risk of holding domestic assets, which in turn, by discouraging capital outflows, would have further enhanced the relationship between capital flows and domestic investment.

Razin and Sadka (2007c) focus on bilateral FDI flows among OECD countries. An important feature of the FDI model (which distinguishes FDI flows from portfolio flows) is fixed setup costs of new investments. This introduces two margins of FDI decisions. There is an intensive margin of determining the magnitude of the flows of FDI, according to standard marginal productivity conditions, and also an extensive margin of determining whether to make a new investment. The first decision gives rise to a flow equation, whereas the second decision produces a selection-condition equation. Crucially, productivity and taxes may affect these two margins in different, possibly conflicting, ways. Yet, unlike our inquiry, Razin and Sadka (2007c) concentrate on FDI flows only and differ in econometric treatment as well as theory.

### 2.2 Gravity in Portfolio Flows

There have been numerous applications of gravity specification in analyzing capital flows. Yet, theoretical foundations are scarce. Martin and Rey (2004) derive a gravity equation for financial holdings when countries trade claims on Arrow Debreu securities. Agents are endowed with a freely traded good, which they can choose to consume, invest in fixed-size risky projects or use to buy shares on the stock market. Investing in a specific project is equivalent to buying an Arrow-Debreu asset that pays in only one state of nature. With endogenously incomplete asset markets economic size becomes important to the determination of asset returns, the breadth of financial markets and the degree of risk sharing and home

bias. Coeurdacier and Martin (2009) analyze trade in bonds, equity and banking assets. The financial version of the gravity equation for the holdings of assets includes a size factor (GDP, the number of assets in source country which relates to economic size (GDP and market capitalization) and the financial sophistication, transaction costs between the two countries, expected returns and the financial price index which is specific to each country. However, these approaches are limited to cases which satisfy Arrow Debreu conditions: when the asset of a country has a positive payoff, the assets of all other countries have a zero payoff.

To avoid this limitation, Okawa and van Wincoop (2012) have recently contributed to the literature on financial flows gravity. The foundations of the model lie in a static portfolio choice framework where investors hold claims on risky assets from a large number of countries. Asset returns are affected both by a country-specific and a global component. Bilateral financial holdings depend on the product of economic size variables (stock market capitalization in the destination country and total investment in stock in the source country) divided by a relative financial friction. The gravity specification for bilateral asset holdings includes the equity supply, equity holdings, World's demand and supply of equity, information frictions, multilateral resistance terms which measure the average financial frictions for a destination and source country.[7]

## 3   Recent Facts

This section briefly overviews some interesting stylized facts on capital flows in developing economies over the last two decades by analyzing aggregate and bilateral flows data. It also reviews some of the factors related to private capital flows commonly discussed in the literature.

### 3.1   Trends in Developing Countries

As shown in Figure 3.1 that describes the evolution of capital flows in low– and middle–income countries, the Financial Crisis has affected both groups considerably. Interestingly, LICs have started to receive substantial amount of capital only very recently (see the size difference in the right scale). The only period of intensified portfolio investments into LICs are a boom period before and, to some extent, after the financial crisis. Even FDI is a very recent phenomenon with a clear upward trend for LICs for a less than a decade.

---

[7] As a limitation, Okawa and van Wincoop (2012) demonstrate that gravity cannot be sustained once one allows for a general covariance structure of asset returns, while assuming that factors generating return co-movement cannot be separately hedged, also for taxes on foreign returns, and trade in only risky assets.

Fig. 3.1: Portfolio Investment in LICs and MICs (right scale: LICs) and FDI in LICs and MICs (right scale: LICs)
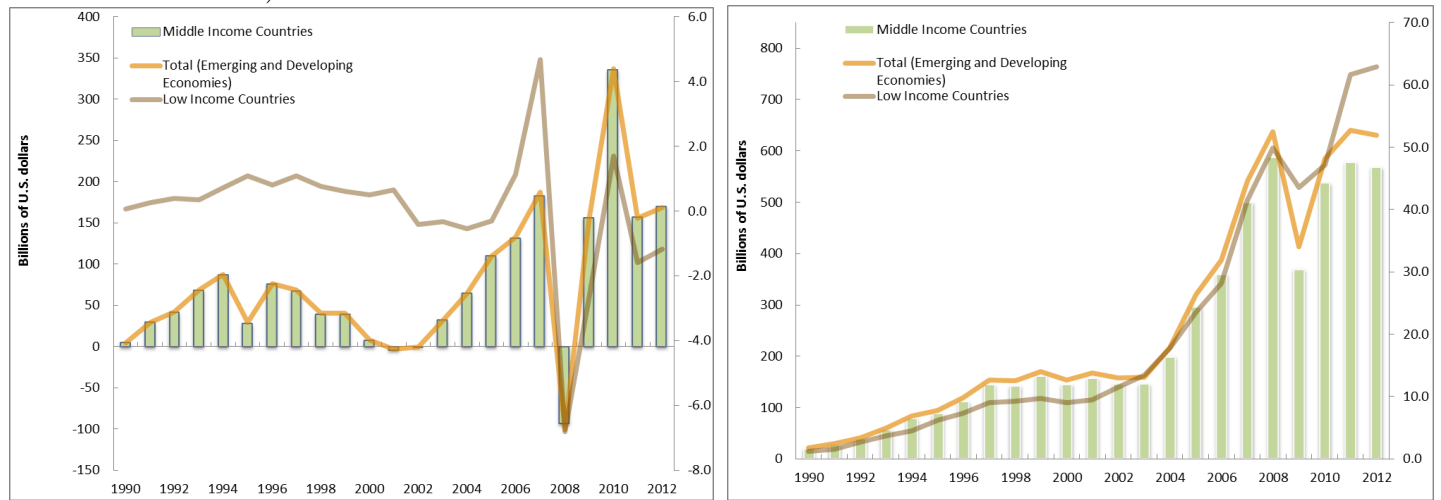


Data: IMF, World Economy Outlook

Though FDI is dominating (see Figure 3.2), portfolio flows, as mentioned, are to be taking some acceleration in recent years too. Due to requirements on some financial development, portfolio flows are more important for a few selected economies rather than are as widely spread out as FDI. Moreover, Figure 3.2 confirms the idea of FDI being relatively stable, most probably because of large fixed costs and thus lower liquidity, as compared to the portfolio flows whose variation on both positive and negative sides is substantial.

Fig. 3.2: Composition of Capital Flows for LICs
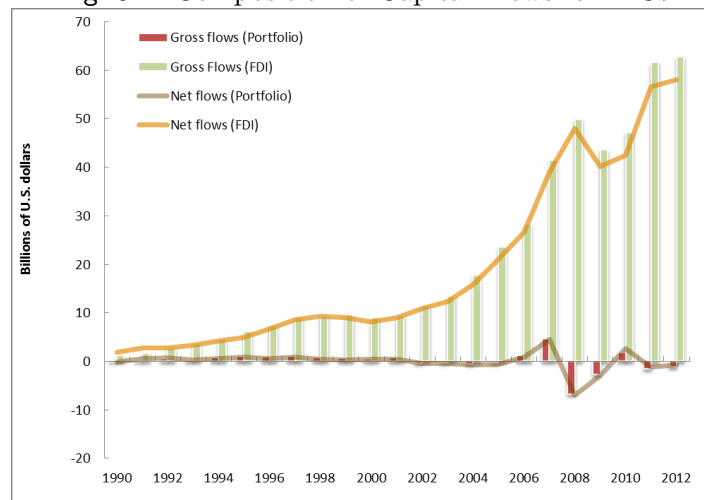


Data: IMF, World Economy Outlook

## 3.2 Intensive and Extensive Margins

In order to further understand the recent development in capital flows in developing countries, we decompose bilateral flows data on extensive and intensive margins. We draw from Felbermayr and Kohler (2006) which distinguishes between changes in the number of active bilateral trade relationships (extensive margin), and the growth of trade volumes in existing relationships (intensive margin). In our context, if attraction in some cases is not strong enough to generate investment at all, then ignoring such cases

altogether implies that we systematically overestimate the force of attraction, or - equivalently - underestimate the investment-inhibiting force of spatial frictions. Push and pull factors are accounted for by bilateral gravity model with both *intensive* and *extensive* margins. The latter allows modeling newly established links, especially among newly emerged markets.

We use a similar approach for the capital flows analysis. There are a few reasons to concentrate on these two margins rather than the aggregate capital flow. At first, investigation of the extensive margin relates directly to diversification issues and welfare. If flows are based on portfolio diversification, then expansion of world financial movements on the extensive margin seems particularly important from a welfare perspective, since it increases the degree of diversification. Hence, we aim at consistently capturing simultaneous movements on both margins through time.

Moreover, there are many zero capital flows, especially to developing countries. Then, fixed costs generate decisions at two margins, i.e., the direction and volume. It also helps establishing the path-dependency observed in investors' behavior (see Welch (1992) for the example with the initial public offering shares and references to many other applications). Further, fixed cost allows addressing the emergence of new destinations for investments once the structure of those costs in the country changes relative to the global system.

Our analysis makes it possible to track which factors are responsible for each of the margins, explore the relative importance of both margins over time, and compare the portfolio and direct investment at each of the margins. Moreover, we can do so for countries at different development stages.

To formalize, let us define $\mathcal{N}_{t,h}$ to denote the number of capital links of vintage $h$ that are active at time $t$. The vintage of a capital link is defined as the earliest time at which capital movements occur between a specific pair of two countries, based on their financial openness. A country is said to be financially open if it receives capital from at least one foreign country. Total capital flow can therefore be written as

$$M_t = \sum_{h=t_0}^{t} \bar{M}_{t,h} \mathcal{N}_{t,h},$$

where $\bar{M}_{t,h}$ is the average bilateral capital flow based on active links of vintage $h$. Further, let $\bar{M}_t$ be the average volume of capital flows at time $t$ across all vintages, $\bar{M}_t = \sum_{h=t_0}^{t} \bar{M}_{t,h} / \sum_{h=t_0}^{t} \mathcal{N}_{t,h}$. Let's start with a change in capital flows on the intensive margin which is defined as the movements in capital volumes over the existing relationships (links):

$$\triangle M_t^{\text{int}} = \sum_{h=t_0}^{t-1} \left( \bar{M}_{t,h} - \bar{M}_{t-1,h} \right) \mathcal{N}_{t-1,h} = \sum_{h=t_0}^{t-1} \triangle \bar{M}_{t,h} \mathcal{N}_{t-1,h},$$

whereas the changes in the number of active links are defined as

$$\triangle N_t = \sum_{h=t_0}^{t-1} \left( \mathcal{N}_{t,h} - \mathcal{N}_{t-1,h} \right) = \sum_{h=t_0}^{t-1} \triangle \mathcal{N}_{t,h}.$$

One can easily check that a change in capital flows is defined as the sum of changes on intensive and extensive margins, $\triangle M_t = \triangle M_t^{\text{int}} + \triangle \mathcal{N}_t \bar{M}_t$,[8]

$$\triangle M_t \equiv \triangle \bar{M}_t \mathcal{N}_{t-1} + \triangle \mathcal{N}_t \bar{M}_t$$
$$= \left( \bar{M}_t - \bar{M}_{t-1} \right) \mathcal{N}_{t-1} + \left( \mathcal{N}_t - \mathcal{N}_{t-1} \right) \bar{M}_t = \mathcal{N}_t \bar{M}_t - \bar{M}_{t-1} \mathcal{N}_{t-1}.$$

Therefore, changes on the extensive margin are "weighted" by end-of-period average flow volumes, while changes on the intensive margin are "weighted" by beginning-of-period numbers of links. Our temporal dimension is short, therefore, results are not to be significantly affected by time weights.

---

[8] Notice that $\triangle M_t^{\text{int}} = \sum_{h=t_0}^{t-1} \triangle \bar{M}_{t,h} \mathcal{N}_{t-1,h} = \triangle \left( \sum_{h=t_0}^{t} \bar{M}_{t,h} / \sum_{h=t_0}^{t} \mathcal{N}_{t,h} \right) \sum_{h=t_0}^{t-1} \mathcal{N}_{t-1,h} = \triangle \bar{M}_t \mathcal{N}_{t-1}.$

Fig. 3.3: Cumulative Difference of Total, Intensive and Extensive Margins of FDI flows for All Developing, Middle–Income and Low–Income Countries from 2002 to 2010
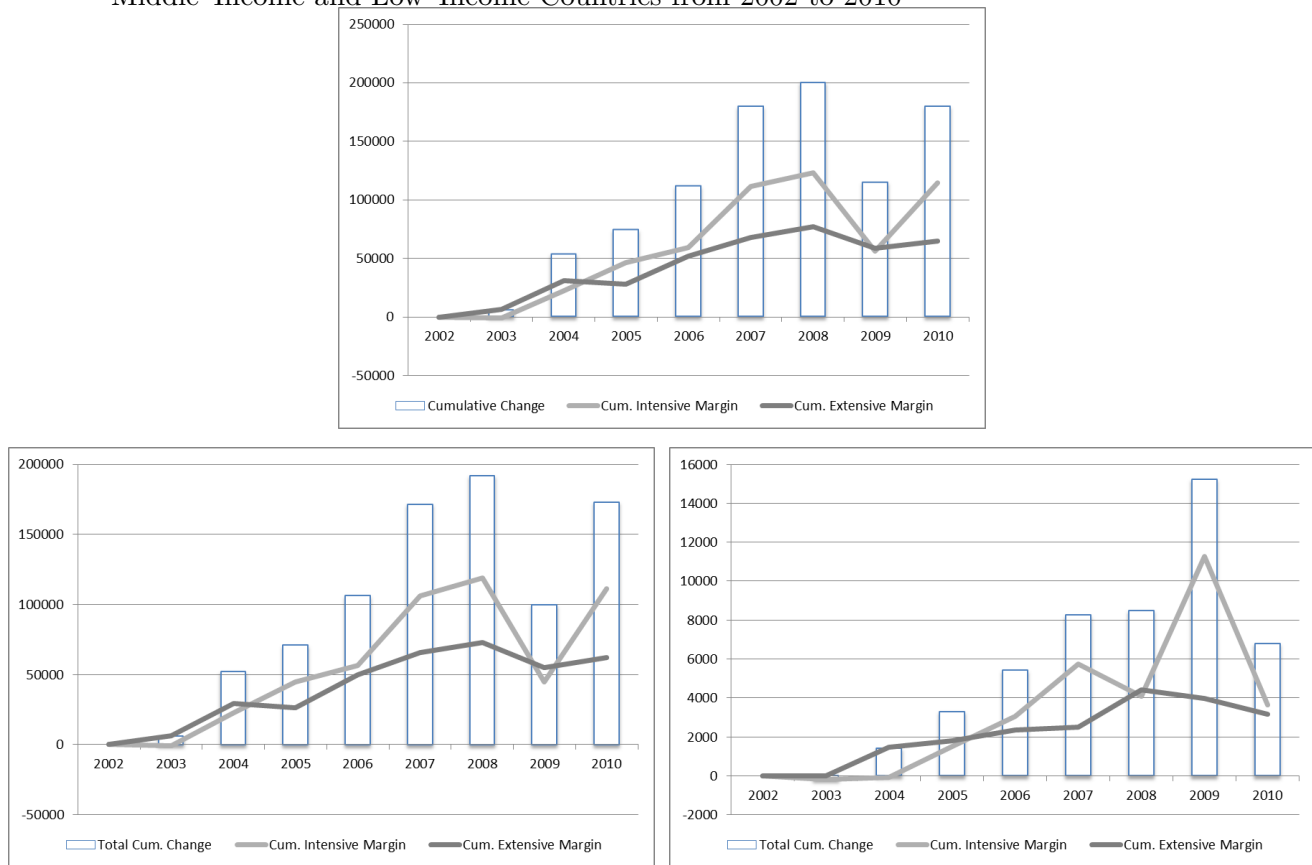


In Figure 3.3 we plot the changes in capital flows over time period 2003-2010 and their decomposition of the intensive and extensive margins. The graph helps exploring changes in domination of the respective margins in good times and over economic crisis along the cross-sectional and the time-series dimension of the data.

The variability of both intensive and extensive margins is substantial for the developing countries. Usually the adjustment takes place at the intensive margin because extensive margin requires reinitiating of the link after it has been broken which may be costly. Indeed, the positive capital flow between two countries arises only if their bilateral asset trade and investment potential exceeds some source-destination economy specific threshold value. Maintaining flow relationships requires certain infrastructure and institutions that facilitate an efficient flow of information and exchange of ideas, in addition to the best-practice banking system. Some relevant institutions are international payment systems, legal agreements, consulates, or the activities of the respective chambers of commerce. Individual firms will not be able to invest in the absence of services provided by such institutions. Rational governments will invest into these institutions only if the expected benefit from a materializing capital flow potential in present value terms exceeds the investment cost. Therefore, capital inflow will only materialize if its full potential exceeds the threshold level, otherwise bilateral flow will be zero.

## 3.3 Expanding Network of Capital Flows

We also analyze the changes in capital flows network. Figure 3.4 demonstrates how capital flows flew from advanced to developing countries in 2002 and 2010. In less than ten years, the network became way denser, with more investors, new receivers and more connections. This calls for further study of capital flow dynamics, in particular, to understand when the link is established and how much is being invested

Fig. 3.4: Network Structure of Capital Flows to LICs in 2002 and 2010
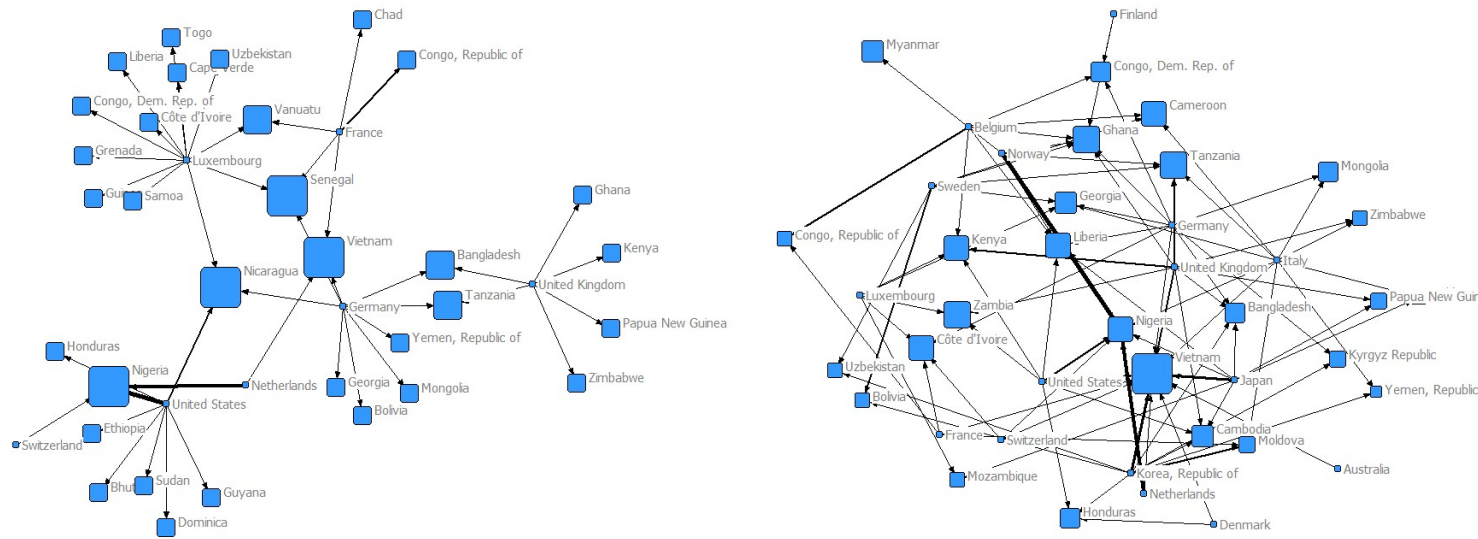


Tab. 3.1: Aggregate Statistics of the Network of Capital Flows in 2002 and 2010 for LICs and MICs

|  | Average Degree | Weighted Average Degree | Core Membership |
|---|---|---|---|
| LIC 2002 | 7.335 | 21.736 | Netherlands, Nigeria, United States |
| LIC 2010 | 9.819 | 103.046 | Nigeria, Netherlands, Norway |
| MIC 2002 | 95.105 | 602.331 | Mexico, United States |
| MIC 2010 | 30.597 | 2600.767 | Brazil, China, India, Japan, Luxembourg, Switzerland, United States |
| Average Degree refers to the sum of edges of a vertex. Weighted Average Degree is the average of sum of weights of the edges of nodes. | | | |

once it is in place. The cross-sectional dependence is also confirmed by a number of statistical tests.[9]

Table 3.1 collects information on aggregate statistics of capital network for both, LICs and MICs. Notice it is exactly for LICs that the average degree has increased - hence, the extensive margin played the most prominent role (the number of links attached to an economy increased). However, both extensive and intensive margins are taken into account once we compute the weighted average degree where the weights (volumes of capital flows) are considered. It is then clear that both LICs and MICs experienced substantial increases in capital flows, with a few economies receiving very large flows. Also, the core members have increased in number and changed identity. The emerging markets - such as Brazil, India and China (BIC) - emerged as core members in 2010 as compared to 2002.

## 3.4 Explanatory Factors

We now turn attention to factors which are particularly relevant for LICs. Our interest in the topic has been sparked by a steep rise in private capital flows to the emerging and developing countries in the middle of the past decade. Although such flows briefly reversed during the apex of the crisis, very low interest rates in advanced countries and an attenuation of global risk aversion have once again prompted investors to scour the globe in search of attractive investment opportunities. The private flows are increasingly the main source of external financing for many countries in the developing world. Moreover, private flows tend to be more volatile in low income countries.

---

[9] Available upon request.

Generally, factors affecting the capital flows can be grouped into global and local factors that influence the supply of funds (for example, including the relative return between the source country and the recipient) and factors that determine the demand for funds. Local economic environment in general and country-specific pull factors in particular are important determinants of the pattern of flows across developing regions. Equity investors are mainly attracted by locations which can offer high, long-term returns and low correlations with other markets.

Among developing countries, the so-called frontier markets[10] managed to attract relatively more capital than other locations. Frontier market short- and medium-term securities typically have higher yields than in more developed emerging countries. Output growth, which supports returns on equity portfolio and foreign direct investment (FDI), has usually been higher in the frontier markets (FMs) than in other regions during the last few years. Another important factor for FDI inflows is an interest in resource extraction activities, also a diversification of the export base. For example, Sub-Saharan African FMs show the fastest average diversification improvements between 1996 and 2009, and they also attract more capital than non-frontier markets, see IMF (2011).[11] Sectors play a significant part in attracting foreign investment: for instance, Angola has experienced unusually large foreign direct investment into its oil extraction industry, peaking at 40 percent of GDP in 2000, IMF (2011). Another example of Mauritius makes a case for a financial platform for investors bound for Asia and sub-Saharan Africa. With a strong regulatory framework in line with international norms and a resilient economy, Mauritius has benefited from favorable investor sentiment, despite lower domestic interest rates. This points to the existence of investments "hubs" among developing countries, too.

As demonstrated by IMF (2011), developing financial markets and integrating them with global markets, and implementing policies that provide a supportive macroeconomic and institutional environment are crucial for capital inflows that can support investment and growth. Yet, all these factors show no structure, channels and investment margins that are affected. We will seek a transparent, parsimonious and more structural approach to major determinants. Though not all determinants can be explicitly modeled or controlled for, the empirical exercise must be flexible to allow for country-level and global factors.

Nonetheless, figure 3.5 demonstrates unconditional within variation of log FDI flows deviations from individual means against log of destination income deviations from individual means. There is a clear upward trend, thereby indicating of gravity forces - destination and source incomes being positively associated with FDI - at work.

---

[10] The term "frontier markets" (FMs) is commonly used to describe a subset of emerging markets (EMs) that have small financial sectors and/or have low annual turnover and liquidity, but nonetheless demonstrate a relative openness to and accessibility for foreign investors. They are generally in the early stages of financial market development. See IMF (2011).

[11] Hence, capital flows are related to trade flows. This diversification offers new opportunities for investors and also requires more capital due to underdevelopment but also potentially offers higher returns to investment. This points to the location and spatial considerations.

Fig. 3.5: Bilateral FDI vs Destination and Source Incomes in Overall Sample
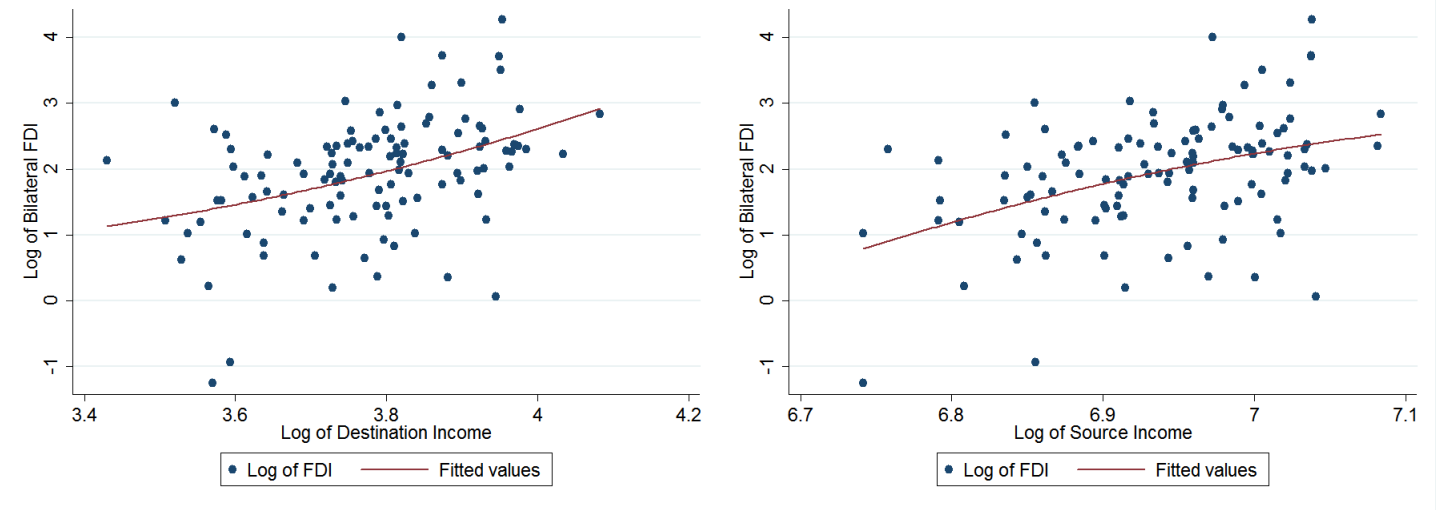


Having documented some important changes in the nature of capital flows in developing economies and especially in LICs, We turn to discussing the modeling strategy which must account for the two margins (intensive and extensive) whose dynamics seem crucial to understanding current and future developments of capital flow changes.

## 4    Theoretical Motivation

Our theoretical setup incorporates decisions to consume, also portfolio investment, and production choices, including investing abroad. Within the same framework, we introduce the FDI flows which require several changes as the driving forces are presumably different compared to portfolio investment. Mainly FDI are flows within multinationals or large companies whose motives can also be described by the gravity equation. For the direct investment, we mainly follow Helpman, Melitz, and Rubinstein (2008) and, instead of trade flows, adapt it to the capital flows.[12] As an extension, we also consider portfolio flows, for which we mainly draw from Okawa and van Wincoop (2012). To produce such a framework, we need to consider both, consumption and production sides of the economy.[13] The simplifying assumption will be separability of total expenditure into (portfolio) investment and goods produced by firms located in different markets (FDI). Firms will be deciding on FDI whereas the residual income will be invested to diversify portfolio by the agents. It is a multi-country setting featuring $\mathbb{N}$ countries indexed by $i = 0, \dots, \mathbb{N}$. Unlike standard models with labor as the only factor (for example, Anderson and van Wincoop (2003), Helpman, Melitz, and Rubinstein (2008), Bergstrand, Egger, and Larch (2008), among others) we assume two factors of production, labor $l$ and capital $k$. Country $i$ is endowed with $L_i$ and $K_i$ aggregate units of inelastically supplied labor and capital, respectively.

Consider a simple two-period problem of agents in country $i$ to choose aggregate consumption ($Q_i$), aggregate investment ($I_i$) and the allocation of consumption and investment goods across $\mathbb{N}$ countries (respectively, $q_{in}$ and $i_{in}$ for all $n$) to maximize the present discounted value of lifetime utility

$$\max U_i \equiv \max_{\{Q_i, I_i, i_i, q_i\} \geq 0} \sum_{t=0}^{T} (1 + \rho)^{-t} U(Q_{it}),$$

subject to

$$P_{it} Q_{it} + P_{it}^{I} I_{it} = E_{it},$$

---

[12] A related piece of work is due to Helpman, Melitz, and Yeaple (2004) which has been a simplified version of firm selection with more symmetry than in Helpman, Melitz, and Rubinstein (2008). Similarly, heterogenous firm framework has been utilized by Kleinert and Toubal (2010) which derive gravity relationship from the three different models of multinational firms.

[13] As was done in Yotov and Olivero (2012) to model dynamic conditional gravity.

where $P$ refers to aggregate price index for goods and $P^I$ stands for the aggregate price for all country $i$ equity claims. Time is discrete and, in principle, $T$ can be infinite, though we limit our analysis to the most trivial dynamics (mainly to reflect a very limited time dimension in empirical exercise). The residual income will be denoted as $\tilde{E}_{it} \equiv E_{it} - P_{it}^I I_{it}$ to refer to the income after the portfolio investment has been subtracted and therefore can be used for consumption, and, similarly, $\mathcal{E}_{it} \equiv E_{it} - P_{it}Q_{it}$. Further, the consumption and investment goods are aggregated by the CES aggregators, i.e.,

$$I_{it} = \left( \int_{\omega \in \Omega_i} i_{it}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}.$$

Hence, the lifetime utility is maximized subject to the budget constraint which accounts for spending on both consumption and investment goods across all countries. The production function of the representative firm in country $i$ producing a brand $\omega$ is assumed to be,

$$y_{it}(\omega) = \frac{k_{it}^{\alpha_i} l_{it}^{1-\alpha_i}}{\varphi},$$

where $y_{it}(\omega)$ summarizes output of variety $\omega$, $l_{it}(\omega)$ is labor used in its production, $k_{it}(\omega)$ is capital used in its production, $\alpha_i \in (0, 1)$ denotes the factor intensity of production in country $i$ and $\varphi$ is a firm's from country $i$'s idiosyncratic productivity innovation drawn from the known distribution $G(\varphi)$. A firm uses $k$ units of capital and $l$ units of labor in variable production.[14] Then the maximizing choice $k_{i,t}(\omega)$ is characterized by the first order condition with respect to aggregate consumption,

$$U_i'(Q_{it}) = \frac{1}{1+\rho} U_i'(Q_{i,t+1}).$$

To simplify our analysis, we choose iso-elastic utility function which helps obtaining closed-form solutions for portfolio allocations, also makes consumption and investment decisions independent of each other.[15] Recall that for constant relative risk aversion utility the investment decision is independent of the consumption decision (Samuelson, 1969 and Merton, 1969). Therefore, agents will be deciding on consumption and portfolio investment separately whereas heterogeneous firms will be deciding on domestic production, exporting and direct investment. We turn to firm choices first, and, for convenience, omit time subscript when it is not essential.

## 4.1 Firms Decisions

A firm with productivity $\varphi$ from country $i$ incurs a cost

$$C_{in}(q) = v_i(q\tau_{in}\varphi + f_{in})$$

to supply $q > 0$ units of goods to country $n$ where $f_{in}$ stands for fixed costs to enter $n$'s market. We refer to $v_i$ as the cost of the composite input bundle in country $i$,

$$v_i(\omega) = \frac{r_i^{\alpha_i} w_i^{1-\alpha_i}}{\alpha_i^{\alpha_i}(1-\alpha_i)^{1-\alpha_i}},$$

---

[14] We abstract from total factor productivity and its stochastic implications - the extension undertaken by Lastauskas and Pesaran (2013) in the gravity modeling with growing economies and global perspective.

[15] Isoelastic functions exhibit constant elasticities. A class of such functions is known as monomials, loosely understood as polynomials with only one term. The functional form is not operational for FDI but is important for demand driven portfolio flows, as covered in Appendix A.2.

where $w_i$ and $r_i$ are country $i$'s payments for labor and capital, respectively.[16] The marginal costs are

$$c_{in}(\varphi) \equiv \frac{dC_{in}(q)}{dq} = v_i \tau_{in} \varphi.$$

Profit maximization implies that a firm charges a constant markup over its marginal cost, $p_{in}(\varphi) = \frac{\sigma}{\sigma-1} c_{in}(\varphi)$. In this case, a firm's market-specific revenue is proportional to its marginal cost,

$$R_{in}(\varphi) = \frac{Q_n P_n}{P_n^{1-\sigma}} \left( \frac{\sigma}{\sigma-1} c_{in}(\varphi) \right)^{1-\sigma} = \tilde{E}_n \left( \frac{p_{in}(\varphi)}{P_n} \right)^{1-\sigma}, \tag{4.1}$$

and its market-specific variable profit is proportional to its revenue $\pi_{in}(\varphi) = R_{in}(\varphi)/\sigma$. We denote expenditure on varieties as the residual expenditure, $\tilde{E}_{nt} \equiv E_{it} - P_{it}^I I_{it}$. With Cobb-Douglas production functions, payments to labor and capital can be expressed as functions of revenue,

$$w_i l_{in}(\varphi) = \frac{\sigma-1}{\sigma}(1-\alpha_i) R_{in}(\varphi),$$
$$r_i k_{in}(\varphi) = \frac{\sigma-1}{\sigma}\alpha_i R_{in}(\varphi).$$

Firms engage in monopolistic competition. Because of Dixit-Stiglitz preferences, monopolists charge a constant markup over marginal cost. Moreover, under monopolistic competition, zero economic profits in equilibrium ensures constant output of each firm. The utility maximization subject to $\tilde{E}_i = \sum_n M_{in}$ yields

$$M_{in} = \left( \frac{p_i \tau_{in}}{P_n} \right)^{-\sigma} \tilde{E}_n \tag{4.2}$$

since income is equal to expenditure in equilibrium. Firms take the residual demand in (4.2) as given and maximize profits by choosing a price. In any country $n$, the price of a variety $\omega$ from country $i$ is given by

$$\begin{aligned} p_{in}(\omega) &= \frac{\sigma}{\sigma-1} \tau_{in} v_i \varphi \\ &= \tau_{in} p_i(\omega) \end{aligned} \quad \text{for all } \omega \in \Omega_{in}. \tag{4.3}$$

Note that $p_{in}$ is the cost-insurance-freight (C.I.F.) price in country $n$ as it depends on trade costs $\tau_{in}$ which could be interpreted as 'iceberg costs'. Hence, for unit to arrive, $\tau_{in}$ more product shall be sent as this share 'melts' on the way. It is clear now that $p_i$ is the free on board (F.O.B.) price locally at exporter.

It is a stylized fact that only a part of firms actually trade, see Bernard, Redding, and Schott (2007). In our setup product and firm partitioning is fully determined by the productivity space since productivities map one-to-one to varieties. Therefore, we consider a measure of firms in country $i$ which supply market $n$. This happens if the variable profit a firm earns there covers its fixed market access cost, $R_{in}(\varphi)/\sigma \geq v_i f_{in}^{EXP}$. Denote by $\varphi_{in}^{EXP}$ the productivity threshold at which the least productive firm from country $i$ sells in country $n$

$$\left( \varphi_{in}^{EXP} \tau_{in} \right)^{1-\sigma} \frac{\tilde{E}_n}{\sigma^\sigma (\sigma-1)^{1-\sigma} P_n^{1-\sigma}} = v_i^\sigma f_{in}^{EXP},$$

whereas the productivity to invest in country $n$ is defined as

$$\left( \varphi_{in}^{FDI} \right)^{1-\sigma} \frac{\tilde{E}_n}{\sigma^\sigma (\sigma-1)^{1-\sigma} P_n^{1-\sigma}} = v_n^\sigma f_{in}^{FDI},$$

noting that no iceberg cost is paid when a firm choose FDI over export and production costs are priced in destination country.[17] To enter, a firm incurs a fixed entry cost of $f^e > 0$ units of the composite

---

[16] We abstract from modelling exchange rate dynamics both theoretically and empirically, mainly because of the data for our targeted, developing, countries. It should be said, however, that Gourinchas and Rey (2007) emphasize that global imbalances and external imbalances (NXA) adjusted via the trade and the valuation channels. Hence, NXA is a linear combination of exports/imports and net foreign asset position. Corte, Sarno, and Sestieri (2012) find a strong predictive power for US dollar rates as well as other numeraire currencies for NXA on a bilateral basis across four major countries.

[17] Note that the condition for ordering of the cutoffs between FDI and exporting is as follows: $\varphi_{in}^{FDI} < \varphi_{in}^{EXP}$ leading to $\varphi_{in}^{EXP} = \frac{1}{\tau_{in}} \left[ \frac{v_i^\sigma f_{in}^{EXP} \sigma^\sigma (\sigma-1)^{1-\sigma} P_n^{1-\sigma}}{\tilde{E}_n} \right]^{1/(1-\sigma)}$ and $\varphi_{in}^{FDI} = \left[ \frac{v_n^\sigma f_{in}^{FDI} \sigma^\sigma (\sigma-1)^{1-\sigma} P_n^{1-\sigma}}{\tilde{E}_n} \right]^{1/(1-\sigma)}$. Therefore, the requirement is equivalent to $\tau_{in}^{\sigma-1} f_{in}^{EXP} < \left( \frac{v_n}{v_i} \right)^\sigma f_{in}^{FDI}$, where the outcome is determined by relative sizes of fixed costs, transportation costs, and relative costs of capital and labor as embodied in $v$.

input bundle $v_i$, making startup costs in country $i$ to be equal to $f^e v_i$. Upon entry, firms draw their productivity $\varphi$ from the same distribution $G(\varphi)$, assumed to be truncated Pareto.[18] It is defined as $G(\varphi) = \frac{\varphi^k - \varphi_L^k}{\varphi_H^k - \varphi_L^k}$, $k > \sigma - 1$. This framework allows for asymmetric capital flows, $M_{ni} \neq M_{in}$, which may also be unidirectional, with $M_{ni} > 0$ and $M_{in} = 0$, or $M_{ni} = 0$ and $M_{in} > 0$.[19] In this setup, the total direct investment from country $i$ to country $n$ is

$$
\begin{aligned}
M_{in}^{FDI} &= \int_{\varphi_L}^{\varphi_{in}^{FDI}} p_{in}^{FDI}(\omega)^{1-\sigma} \tilde{E}_n P_n^{\sigma-1} N_i dG(\varphi) = \tilde{E}_n P_n^{\sigma-1} N_i \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} v_n^{1-\sigma} \int_{\varphi_L}^{\varphi_{in}^{FDI}} \varphi^{1-\sigma} dG(\varphi) \\
&= \tilde{E}_n P_n^{\sigma-1} N_i \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} v_n^{1-\sigma} \frac{k(\varphi_L)^{k-\sigma+1}}{(\varphi_H^k - \varphi_L^k)(k-\sigma+1)} \left[\left(\frac{\varphi_{in}^{FDI}}{\varphi_L}\right)^{k-\sigma+1} - 1\right] \\
&= \tilde{E}_n P_n^{\sigma-1} N_i \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} v_n^{1-\sigma} V_{in}^{FDI},
\end{aligned}
\tag{4.4}
$$

where $N_i$ is the mass of entering firms which is an endogenous object. Let selection be subsumed within

$$
V_{in}^{FDI} = \frac{k(\varphi_L)^{k-\sigma+1}}{(\varphi_H^k - \varphi_L^k)(k-\sigma+1)} \left[\left(\varphi_{in}^{FDI}/\varphi_L\right)^{k-\sigma+1} - 1\right],
\tag{4.5}
$$

then the price index with firms investing abroad is defined as

$$
\begin{aligned}
P_i^{1-\sigma} &= \sum_{n=0}^{\mathbb{N}} N_n \left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1} v_n^{1-\sigma} \int_{\varphi_L}^{\varphi_{in}} \varphi^{1-\sigma} dG(\varphi) \\
&= \left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1} \sum_{n=0}^{\mathbb{N}} v_n^{1-\sigma} N_n V_{in}^{FDI}.
\end{aligned}
\tag{4.6}
$$

Finally, full employment of production factors is obtained whenever

$$
\begin{aligned}
\sum_{n=1}^{\mathbb{N}} \sum_j w_i l_{in}(\varphi) &= \sum_j \left((1-\alpha_i) \frac{\sigma-1}{\sigma} \sum_{n=1}^{\mathbb{N}} R_{in}(\varphi) + K_i v_i f^e\right), \\
\sum_{n=1}^{\mathbb{N}} \sum_j r_i k_{in}(\varphi) &= \sum_j \left(\alpha_i \frac{\sigma-1}{\sigma} \sum_{n=1}^{\mathbb{N}} R_{in}(\varphi) + K_i v_i f^e\right).
\end{aligned}
$$

In the endogenous entry case, the free entry condition implies that total profits are equal to zero, which leads to $\tilde{E}_i = Q_i P_i = w_i L_i + r_i K_i$ (also constant returns to scale technology if factors are paid their marginal products). Using the result on the number of entering firms and gravity equation

$$
\sum_{n=1}^{\mathbb{N}} M_{in}^{FDI} = \tilde{E}_i = K_i \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \sum_{n=1}^{\mathbb{N}} \tilde{E}_n P_n^{\sigma-1} v_n^{1-\sigma} V_{in}^{FDI},
$$

and substituting for $K_i$ we obtain

$$
\begin{aligned}
M_{in}^{FDI} &= \frac{\tilde{E}_i \tilde{E}_n P_n^{\sigma-1} v_n^{1-\sigma} V_{in}^{FDI}}{\sum_{n=1}^{\mathbb{N}} \tilde{E}_n P_n^{\sigma-1} v_n^{1-\sigma} V_{in}^{FDI}} = \frac{\tilde{E}_i \tilde{E}_n}{\sum_{n=1}^{\mathbb{N}} \tilde{E}_n V_{in}^{FDI} (v_n/P_n)^{1-\sigma}} \left(\frac{v_n}{P_n}\right)^{1-\sigma} V_{in}^{FDI} \\
&= v_n^{1-\sigma} V_{in}^{FDI} \frac{\tilde{E}_i \tilde{E}_n}{[P_n \Pi_i]^{1-\sigma}},
\end{aligned}
\tag{4.7}
$$

---

[18] As evidenced by Aoyama, Fujiwara, and Ikeda (2010), a number of phenomena obey the power-law distribution. It is shown that income and net profit of large Japanese companies follow Pareto law. Moreover, despite the choice of unit, i.e., flows or stocks, Pareto distribution is obtained in both cases for European data. Total capital and sales of French companies and the number of employees of British companies all fit into Pareto framework. Axtell (2001) also finds that Pareto reasonably approximates the observed distribution of firm sizes in the USA for data from multiple years and for various definitions of firm size. As a caveat, however, we note that power laws are less successful in dealing with the smaller establishments, also with predicting a number of links for the most central firms in the network and firms with low in-degrees. See Atalay, Hortacsu, Roberts, and Syverson (2011).

[19] Following Helpman, Melitz, and Rubinstein (2008), we can introduce a latent variable, $Z_{in}^{FDI} \equiv \left(\frac{\varphi_{in}^{FDI}}{\varphi_L}\right)^{\sigma-1}$ which can also be rewritten as $Z_{in}^{FDI} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma} \tilde{E}_n \left(\frac{c_{in}(\varphi_L)}{P_n}\right)^{1-\sigma} / v_n f_{in}^{FDI}$, which is the ratio of variable profits for the most productive firm to the fixed costs for exports from $i$ to $n$. Positive multinational sales are observed if and only if $Z_{in}^{FDI} > 1$. Moreover, scaling $z_{in} \equiv \ln Z_{in}^{FDI}$ by the standard deviation $\sigma_\eta$, we obtain the probability $\rho_{ji}$ that $i$ decides invest to $j$, conditional on the observed variables, $\rho_{ni} = \Pr(z_{ni}^\star > 0 \mid \text{observed variables})$.

where the traditional multilateral resistance terms are included in the gravity,

$$\Pi_i^{1-\sigma} = \sum_{n=1}^{\mathbb{N}} \tilde{E}_n V_{in}^{FDI} \left(v_n/P_n\right)^{1-\sigma}, \tag{4.8}$$

and price index in (4.6) leads to an expression similar to that in Anderson and van Wincoop (2003). The difference of the gravity equation as compared to the traditional expression includes the destination cost term $v_n$ which is an aggregator of the labour costs and capital rental rate. The term $V_{in}^{FDI}$ captures the probability of entering into $n$'s economy and the extensive margin of capital network. Due to separability, $\tilde{E}$ denotes the residual expenditure after the portfolio investments have been subtracted. It is then clear that the flow of capital decreases, ceteris paribus, in destination's wages and cost of capital (recall that $\sigma > 1$), increases in own and destination expenditures, and increases in the multilateral resistance terms which capture all the weighted costs to invest in all other economies.

There are few ways of dealing with $P_n\Pi_i$, the (directional) fixed effects being the most common approach, see Feenstra (2004). Another method is to approximate resistance terms around the center of a free trade world allowing inward and outward terms to be unequal in magnitude. Koch and LeSage (2009) proceed along these lines and allow for errors due to linearization and interdependence between the multilateral resistance indices to be captured by latent unobservable terms consistent with the implicit forms that arise from theory. They translate the global interdependence structure into a Bayesian and produce estimates of the latent unobservable multilateral resistance terms. Behar and Nelson (2009) use the first-order Taylor expansion around a symmetric center in which all countries trade, following Baier and Bergstrand (2009). As is demonstrated by Behar and Nelson (2009), the latent variable, describing whether investment is positive, can be expressed as a function of the multilateral resistance terms. Therefore, the system-wide dependence is embedded in the first stage of the estimation, too.[20] An interested reader is referred to the Appendix A.2 where agents' decisions are covered in detail. They are operational for the portfolio flows, and constitute an extension of the benchmark framework. We will refer to the relevant derivations when cover the empirical results.

## 5   Empirical Analysis

Extensive margin requires to account for the decision to invest in a location out of many existing possibilities. There are several alternatives to model this decision making. To account for both margins, trade literature has used probit (Helpman, Melitz, and Rubinstein, 2008), logit (Crozet and Koenig, 2010), tobit (Felbermayr and Kohler, 2006) and Poisson (Santos-Silva and Tenreyro, 2007, Silva and Tenreyro, 2009) specifications. Though all these models are similar in many ways, they possess some differences. Poisson specification is best suited for modeling data with heteroskedastic disturbances. Probit works best when a large number of dependent variable assumes the value of zero. Probit and tobit differ, however, in how the unobserved flow is translated into the observed one. In the tobit model, we know the value of latent flow $M_{in}^{\star}$ when $M_{in}^{\star} > 0$, while in the probit model we only know if $M_{in}^{\star} > 0$. Since there is more information in the tobit model, the estimates should be more efficient.[21] Henceforth we will concentrate on two

---

[20] See Footnote 19 for the definition of the latent variable $Z_{in}^{FDI}$. In our case, $Z_{in}^{FDI} = \frac{\frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\tilde{E}_n(v_n\varphi_L)^{1-\sigma}P_n^{\sigma-1}}{v_n f_{in}^{FDI}} = \frac{\tilde{E}_n \tilde{E}_i \varphi_L^{1-\sigma}}{\sigma K_i v_n^\sigma f_{in}^{FDI}} (P_n\Pi_i)^{\sigma-1} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\frac{P_n^{\sigma-1}\tilde{E}_n\varphi_L^{1-\sigma}}{v_n^\sigma f_{in}^{FDI}}$. Note that separation of input cost does not work here, as opposed to Behar and Nelson (2009), because FDI relies on destination country's labor hiring and capital rental rates. Econometrically, the extensive margin in our setting should be driven by destination variables only, i.e., destination prices, destination income, destination cost of labor and capital, and fixed costs of direct investment.

[21] Yet, the suitability of the tobit model is testable by multiplying the probit coefficients by $\sigma_{tobit}$ - this should produce close estimates of the tobit coefficients. This is because the probit estimates should be consistent for $\sigma_{tobit}$-adjusted tobit estimates given tobit is the correct model (Greene, 2003). From technical point of view, the tobit loglikelihood function is really a probit model combined with a truncated regression model, with the coefficient vectors in the two models restricted to be proportional to each other. This restriction can be tested by means of an LR test.

most popular methodologies - probit and tobit modeling. We will emphasize both statistical aspects as well as differences in economic treatment of capital flows of value zero. Our novelty includes the explicit treatment of cross-sectional dependence and arbitrary correlation between two, intensive and extensive, margins. The extensive margin is particularly acute for the developing countries as demonstrated in the Section 3 on empirical facts.

## 5.1  Methodology

In a system with many zero capital flows, care must be taken to model zero observations and their relationship with the volumes of capital flows. Note that higher number of investors implies higher asset trade volumes. However, this extensive margin is negatively correlated the financial, institutional and informational frictions, and omitting a control for it would lead to overestimation of these barriers on capital flows. Another source of bias is the selection when country pairs with zero capital flows are excluded. It induces a positive correlation between the unobserved disturbance terms and various flow frictions. We refer to Technical Appendix B for an extensive treatment of probit, pseudo-Poisson maximum likelihood, correlated tobit and panel extensions tailored specifically for our theoretical framework. This way we aim at structural mapping of theory into empirical exercise.

## 5.2  From Theory to Data

In our estimation, we combine a number of datasets, covering 71 destination economies (low and middle income countries) and 25 OECD economies which are investors. The time frame is 2002 to 2010, thus forming a short unbalanced panel. We use IMF (CPIS), OECD (Globalization statistics), World Bank, International Country Risk Guide, and others. We overview the coverage and data issues in Appendix C.

Conversion of theoretical model into an estimable model is not straightforward. We exploit the cost side which has been emphasized in a number of studies. Recall that prices, and thus capital flows in gravity form, are functions of $v_i$, the cost of the composite input bundle. We measure labor productivity using wage data - the two are the same under perfect labor market. Wages are purported to be economically relevant in Egger and Radulescu (2011) too:[22] a firm cares about employer-borne taxes of high skilled workers who are elemental to firm set-up because higher wage costs lead to higher costs of production. Authors employ the average wage in the manufacturing sector; unfortunately, this measure is unavailable for many developing countries in our sample. Therefore, we will use a more aggregate measure.

Razin, Sadka, and Tong (2008) emphasize the role of labor productivity rather than costs in the host country to increase the volume of the FDI flows through the standard marginal profitability effect. However, such a shock may lower the likelihood of making any new FDI flows by the source country through a total profitability effect due to an increase in domestic input prices. A sample of 62 OECD and non-OECD countries over the period 1987-2000 was analyzed using the Heckman procedure with two margins: one decides how much to invest abroad, while ignoring the fixed setup cost; then, a decision is made whether to invest at all, taking into account this cost. The key explanatory variable, productivity is the output per worker as measured by PPP-adjusted real GDP per worker. Productivity is also proxied by fitting a regression on capital over labor ratio and years of schooling. Though theoretically acknowledged,

---

[22] Egger and Radulescu (2011) concentrate on the effective labor tax rates and bilateral outward FDI stocks among 52 countries in 2002 and document that the employee-borne part of labor taxes determines bilateral FDI significantly different from zero. Moreover, personal income tax rates turn out somewhat less important than profit tax rates. Taxes affect FDI in more complicated ways, as emphasized by Razin, Rubinstein, and Sadka (2005), which argue that international tax differentials act on the direction and magnitude of FDI flows in a different way: the source country tax rate works primarily on the selection process, whereas the host-country tax rate affect mainly the magnitude of the FDI, once they occur. In a similar spirit of our analysis, Razin and Sadka (2007c) analyze FDI flows along two margins, intensive and extensive. Their question is the effect of corporate tax rate on FDI flows for 18 OECD economies in a period of 1987-2003. Authors also analyze the role of labor productivity on the flows (the productivity is measured as output per worker in PPP value). Their prediction of a positive effect on the intensive margin with an ambiguous effect on the extensive margin finds some empirical support.

startup costs are not included in estimating equations. A dummy for a positive lagged FDI is used in the selection equation. Its positive coefficient is then interpreted as evidence of a lower threshold barrier for pairs of countries that had positive FDI flows in the past. We allow for dynamics in our extension, however, wanting a more structural approach and already having accounted for the effects of GDP in our estimating equations, we measure labor productivity by using wage data.

## 5.3 Estimating Equilibrium Relations

We seek to map our parsimonious theoretical account of the main forces of capital flows into estimating relations. Our main equations of interest are as follows (the first one refers to (4.7) whereas the second is derived in Appendix, see (A.5)),

$$M_{ni}^{FDI} = v_n^{1-\sigma} V_{in}^{FDI} \frac{\tilde{E}_i \tilde{E}_n}{[P_n \Pi_i]^{1-\sigma}},$$
$$M_{ni}^{I} = \gamma_{ni}(\tau_{ni}) \frac{\mathcal{E}_n \mathcal{E}_i}{W \Pi_n \mathcal{P}_i} \delta_{in},$$

where $\tilde{E}_i = \sum_{n=1}^{N} M_{in}^{FDI}$, and, due to separability, $\tilde{E}_n = E_n - \mathcal{E}_n$. In other words, entire expenditure is composed of FDI and portfolio capital. Let us re-express all the equations in estimating equilibrium relations. For that we need to introduce a relative share of portfolio capital in each economy, namely $\omega_i = \mathcal{E}_i / E_i$. In such a case, the main estimating equation (4.7) can be re-expressed as

$$M_{ni}^{FDI} = (1 - \omega_i - \omega_n + \omega_i \omega_n) v_n^{1-\sigma} V_{in}^{FDI} \frac{E_i E_n}{[P_n \Pi_i]^{1-\sigma}}, \tag{5.1}$$

and the portfolio investment as

$$M_{ni}^{I} = \omega_n \omega_i \gamma_{ni}(\tau_{ni}) \frac{E_i E_n}{W \Pi_n \mathcal{P}_i} \delta_{in}.$$

The sum of both types can be written as a function of source and destination income, and a nonlinear term that involves composite costs, informational and other frictions, multilateral and bilateral costs,

$$M_{ni}^{FDI} + M_{ni}^{I} = \left[ \left( \frac{1 - \omega_i - \omega_n + \omega_n \omega_i}{\omega_n \omega_i} \right) \frac{v_n^{1-\sigma} V_{in}^{FDI}}{[P_n \Pi_i]^{1-\sigma}} + \frac{\gamma_{ni}(\tau_{ni}) \delta_{in}}{W \Pi_n \mathcal{P}_i} \right] \omega_n \omega_i E_i E_n. \tag{5.2}$$

Notice that the relative size of portfolio to FDI flows is being pinned down by three intuitive terms, i.e.,

$$\frac{M_{ni}^{I}}{M_{ni}^{FDI}} = \frac{\omega_n \omega_i}{1 - \omega_i - \omega_n + \omega_i \omega_n} \frac{\gamma_{ni}(\tau_{ni})}{v_n^{1-\sigma}} \frac{[P_n \Pi_i]^{1-\sigma}}{W \Pi_n \mathcal{P}_i} \frac{\delta_{in}}{V_{in}^{FDI}}, \tag{5.3}$$

where $\gamma_{ni}(\tau_{ni}) / v_n^{1-\sigma}$ refers to relative (bilateral) variable costs, associated with investment, $(P_n \Pi_i)^{1-\sigma} / (\Pi_n \mathcal{P}_i)$ stands for the relative multilateral resistance terms, and $\delta_{in} / V_{in}^{FDI}$ denotes a ratio of extensive margin terms (factors, affecting the very probability to invest). Taking logs in (5.1) and introducing stochastic shocks, lead to

$$\ln M_{ni}^{FDI} = \beta_{in}^{FDI} + (1 - \sigma)[\alpha_i \ln r_i + (1 - \alpha_i) \ln w_i - \ln P_n - \ln \Pi_i] + \ln E_i + \ln E_n + \ln V_{in}^{FDI} + \varepsilon_{ni}^{FDI},$$
$$\ln M_{ni}^{I} = \beta_{in}^{I} + [\ln \gamma_{ni}(\tau_{ni}) - \ln \mathcal{P}_i - \ln \Pi_n] + \ln E_i + \ln E_n + \ln \delta_{in} + \varepsilon_{ni}^{I},$$
$$\ln \left( M_{ni}^{I} / M_{ni}^{FDI} \right) = \beta_{in}^{I/FDI} + \ln \gamma_{ni}(\tau_{ni}) - (1 - \sigma)[\alpha_i \ln r_i + (1 - \alpha_i) \ln w_i] + \left[ \ln \delta_{in} - \ln V_{in}^{FDI} \right], \tag{5.4}$$

which justifies using the same expenditure (income) aggregates in both, FDI and portfolio flows.[23] We will use startup costs to model the decision to invest (FDI) and contract enforcement (FPI).[24] We proceed

---

[23] The fixed source-destination effects are subsumed within and $\beta_{in}^{FDI} \equiv \ln \left( \frac{1 - \omega_i - \omega_n + \omega_i \omega_n}{(1-\sigma)\left(\alpha_i^{\alpha_i}(1-\alpha_i)^{1-\alpha_i}\right)} \right)$, $\beta_{in}^{FDI} \equiv \ln \left( \frac{\omega_n \omega_i}{W} \right)$ and $\beta_{in}^{I/FDI} \equiv \ln \left( \frac{\omega_n \omega_i (1-\sigma)\left(\alpha_i^{\alpha_i}(1-\alpha_i)^{1-\alpha_i}\right)}{1 - \omega_i - \omega_n + \omega_i \omega_n} \frac{[P_n \Pi_i]^{1-\sigma}}{W \Pi_n \mathcal{P}_i} \right)$.

[24] Notice that contract enforcement will be used to proxy for both, $\gamma_{ni}(\tau_{ni})$ and $\delta_{in}$. We will also use startup costs to proxy for institutional setting, also for the purposes to see whether portfolio flow can be treated as the substitute when costs establishing a long-run business relationship are too high.

Tab. 6.1: Panel Results of Bilateral FDI Ignoring the Extensive Margin

| Variables | LIC | | | MIC | | | Overall Sample | | |
|---|---|---|---|---|---|---|---|---|---|
| | Pooled OLS | FE | RE (GLS) | Pooled OLS | FE | RE (GLS) | Pooled OLS | FE | RE (GLS) |
| Log Destination | -1.08 | -0.80** | -1.07 | -1.28** | -1.28*** | -1.28*** | -1.18** | -1.14*** | -1.18*** |
| Income | (0.852) | (0.344) | (0.976) | (0.564) | (0.467) | (0.482) | (0.477) | (0.418) | (0.430) |
| Log Source | 3.66*** | 4.14*** | 3.68*** | 5.26*** | 5.32*** | 5.26*** | 4.99*** | 5.080*** | 5.00*** |
| Income | (1.153) | (0.958) | (1.120) | (0.744) | (0.744) | (0.578) | (0.634) | (0.638) | (0.512) |
| Log Lending | -1.11*** | -1.03*** | -1.10* | -0.23** | -0.27*** | -0.24** | -0.30*** | -0.32*** | -0.30*** |
| Rate | (0.365) | (0.157) | (0.349) | (0.110) | (0.090) | (0.099) | (0.106) | (0.071) | (0.096) |
| Log Labor | 3.36*** | 2.38*** | 3.32** | 2.51*** | 2.56*** | 2.51*** | 2.51*** | 2.48*** | 2.51*** |
| Productivity | (1.288) | (1.005) | (1.360) | (0.640) | (0.508) | (0.549) | (0.574) | (0.416) | (0.504) |
| Observations | 736 | 736 | 736 | 3209 | 3209 | 3209 | 3945 | 3945 | 3945 |
| Within $R^2$ | | 0.15 | 0.15 | | 0.20 | 0.20 | | 0.19 | 0.19 |
| Between $R^2$ | | | 0.56 | | | 0.64 | | | 0.64 |
| Overall $R^2$ | | | 0.47 | | | 0.56 | | | 0.57 |

Note: time coverage is 2002-2010. Pooled OLS refer to pooled least squares where any time-specific effects are assumed to be fixed and the individual affects are centered around a common intercept; cluster robust standard errors are reported, also directional fixed effects (source and destination countries) are included. FE stands for fixed effects with the reported Driscoll-Kraay standard errors which help controlling for spatial (cross-sectional) dependence. RE (GLS) for random effects generalized least squares estimator (effectively a weighted average of within and between estimators which are not reported separately), also directional fixed effects (source and destination countries) are included. Robust Hausman test has been conducted which works under clustered standard errors therefore not making a claim of RE being efficient (see Wooldridge (2010) for the procedure). Fixed effects are preferred. Standard errors in parentheses; *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

under assumption $\varepsilon_{ni}^{FDI} \perp \varepsilon_{ni}^{I}$ which can be relaxed estimating both flows simultaneously with unspecified covariance structure. The two equations are functions of income at investor and the destination economy, global effects are captured by the multilateral resistance terms, fixed effects are subsumed within the income shares $\omega$, and selection effects are taken care by $V_{in}^{FDI}$ and $\delta_{in}$ (see equations (4.5) and Appendix (A.4)).

Cragg's original specification was relaxed by Jones (1992) to allow for correlated errors such that

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\rho \\ \sigma_2\rho & 1 \end{pmatrix}.$$

Notice that both parts of the (normal) likelihood function (where the lower case variables denote logs of original ones, i.e., $\mathsf{m}_{in} \equiv \ln M_{in}$),

$$l_i(\theta) = 1[M_{in} = 0]\ln[1 - \Phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})]$$
$$+1[M_{in} > 0]\left\{\ln\left[\Phi\left([\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta)](1-\rho^2)^{-\frac{1}{2}}\right)\right] + \ln[\phi((\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta)/\sigma)] - \ln\sigma - \mathsf{m}_{in}\right\},$$

must, however, be maximized simultaneously as there is no two-step procedure to split the likelihood. Literature has emphasized that if the assumption of homoskedastic, normally-distributed, errors is violated, then ML estimates are inconsistent.[25] Moreover, idiosyncratic shocks for the two margins are likely to be correlated, thus causing efficiency issues.

## 6  Baseline Results

We first consider results when the two margins are modelled separately. Table 6.1 demonstrates the results had we fully ignored the adjustment at extensive margin. Results do not account for changes

---

[25] See Robinson (1982) which demonstrates that ML estimation of latent dummy variable models yields inconsistent estimates if the assumption of normality does not hold.

Tab. 6.2: Panel Results of Bilateral Portfolio Ignoring the Extensive Margin

| | Overall Sample | | |
| --- | --- | --- | --- |
| Variables | Pooled OLS | FE | RE (GLS) |
| Log Destination | 0.269*** | 0.0866*** | 0.0837*** |
| Income | (0.0185) | (0.0266) | (0.0259) |
| Log Source | 0.1567*** | 1.083*** | 1.032*** |
| Income | (0.0293) | (0.0662) | (0.0634) |
| Contract Risk | 0.496** | 0.425* | 0.476** |
| | (0.213) | (0.236) | (0.231) |
| Observations | 2,738 | 2,738 | 2,738 |
| Within $R^2$ | | 0.190 | 0.1897 |
| Between $R^2$ | | | 0.7773 |
| Overall $R^2$ | | | 0.6659 |

Note: time coverage is 2002-2010. Pooled OLS refer to pooled least squares where any time-specific effects are assumed to be fixed and the individual affects are centered around a common intercept; cluster robust standard errors are reported, also directional fixed effects (source and destination countries) are included. FE stands for fixed effects with the reported Driscoll-Kraay standard errors which help controlling for spatial (cross-sectional) dependence. RE (GLS) for random effects generalized least squares estimator (effectively a weighted average of within and between estimators which are not reported separately), also directional fixed effects (source and destination countries) are included. Standard errors in parentheses; *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

in capital network evolution, effectively treating the variable $V_{in}^{FDI}$ as irrelevant. The measures of fit refer to within (R-squared from the mean-deviated regression), between (the squared correlation between predicted values using FE and the within-individual means of the original dependent variable) and overall (the squared correlation between predicted values using FE and the untransformed dependent variable). For completeness, we report the results using marginal product of capital rather than the lending rate in the Appendix D.1. As neoclassical theory suggests, the lower income abroad should incentivize the investment - this seems to be the case for MICs and overall sample (the elasticity of destination income is negative). Investment increases in source income and destination labor productivity (wages) but decreases in capital rental rate (lending rate).

Similarly, Table 6.2 reports results of intensive margin only for portfolio flows (see estimating equations reported in (5.4) which underlie the following results). The surprising result concerns contract variable whose higher value indicates higher risk of contract violation. This might be the reflection of the country group included in the analysis which are being targets because of higher returns despite excess risk associates with capital expropriation. Alternatively, this can illustrate the misspecification using the intensive margin only.

Another extreme is to consider extensive margin only. Though our largely static theory does not directly lead to accounting for the persistence in variables, it is suspected that margins are auto-correlated. Provided fixed costs constituted a substantial part of total costs, we should observe a persistence in extensive margin. In other words, conditioning on the previous period's existence of an investment relationship, should affect decision for the current investment. Therefore, we use a dynamic probit to explore the temporal effects. If the margin at time zero is correlated with unobserved heterogeneity, inconsistent estimators are obtained (an example of the so-called initial conditions problem). Heckman's method is considered to be most precise and beats the existing alternatives (see a study by Miranda (2007) which compares estimators in a Monte Carlo study). In the Table 6.3, we report the first stages (probit) estimates after we allow for the investment probability to depend on the decision in the previous period. As opposed to Table 6.1 and Table 6.2, here the extensive (not the intensive) margin is considered. The results indicate that persistence is found to be quite significant and robust across country groups.[26] However, having a short unbalanced panel, we stick to less computationally intensive methods whose

---

[26] For the methodological issues on how to estimate a dynamic probit model, see Stewart (2006, 2007).

Tab. 6.3: Dynamic Probit Model for the FDI Decision

| Variables | LIC | | MIC | | Overall Sample | |
|---|---|---|---|---|---|---|
| | No DFE | With DFE | No DFE | With DFE | No DFE | With DFE |
| Lagged Dummy for FDI | 1.693*** | 0.958*** | 1.250*** | 0.632*** | 1.439*** | 0.752*** |
| | (0.0677) | (0.0807) | (0.0405) | (0.0482) | (0.0341) | (0.0409) |
| Log Lending Rate | -0.144** | 0.145 | -0.0602 | -0.0722 | -0.105*** | -0.0667 |
| | (0.0697) | (0.284) | (0.0407) | (0.0996) | (0.0339) | (0.0940) |
| Log Labor productivity | -0.0416 | 0.468 | -0.195*** | 0.783*** | 0.111*** | 0.528** |
| | (0.0532) | (0.534) | (0.0453) | (0.285) | (0.0265) | (0.244) |
| Log Startup Costs | -0.0362 | -0.116 | -0.0489** | 0.128* | -0.0450*** | 0.0350 |
| | (0.0315) | (0.0873) | (0.0209) | (0.0665) | (0.0169) | (0.0514) |
| Observations | 2266 | | 4609 | | 6875 | |
| Log likelihood | -971.63 | -776.92 | -2582.19 | -2159.61 | -3638.61 | -2982.78 |
| Pseudo $R^2$ | 0.27 | 0.36 | 0.17 | 0.31 | 0.2361 | 0.37 |

Note: time coverage is 2002-2010. DFE refers to the directional (source and destination) fixed effects. The estimation method is random effects dynamic probit model, estimated by the maximum likelihood. Integrals are evaluated using Gaussian–Hermite quadrature points. Standard errors in parentheses; *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

results are testable in a more straightforward manner and which correspond to the static theory more closely.[27].We interpret the finding of persistence as suggesting that empirically fixed and startup costs need to be included in the static (steady state) specifications for the decision to invest (extensive margin).

Having considered the two margins separately, we map the theory to data by introducing the mechanism that governs not only the choice of capital intensity but also the likelihood of investment. The use of two-part (also known as hurdle) model seems natural as it is mainly driven by economic arguments: the decision to start investing in an economy and the volume of investment are governed by non-identical mechanisms as portrayed in theory. The first part of the model estimates the probability of observing a positive capital flow (known as the selection or participation equation) while the second estimates the volume of capital conditional on observing positive flows. Therefore, such statistical specifications as tobit are too restrictive as they assume a single mechanism for both intensive and extensive margins. Yet, the two-part models which assume conditional independence of deciding on investing and the amount also raises a question of reliability. Therefore, we will also explore models which explicitly allow correlation between two investment margins. Note that the first part of the model estimates the probability of observing a positive capital flow while the second estimates the volume of capital conditional on observing positive flows.[28] Table 6.4 collects result from different methods which account for the two decisions. Lognormal hurdle refers to the model where the variable can be decomposed into a product of an indicator function and lognormally distributed capital flow, provided it is positive.

The panel data two-part models comply well to the traditional pooling estimation absent unobserved effects (see Wooldridge, 2010). The needed adaptation is the robust variance matrix estimator to account for the serial correlation in the score across time dimension.[29] However, the existence of unobserved effects creates an incidental parameters problem for "small" $T$ panels when we want to estimate these effects together with the other model's parameters and error variances. One of the solutions is to assume independence of the unobserved effects and covariates along with several other assumptions. This would

---

[27] Not only initial condition problem needs to be addressed but also multivariate normal probability functions of order $T$ are required. One possibility is to resort to simulated likelihood method and estimate a simulated multivariate random effects probit. Relying on underlying assumptions and imposed structure, we feel a static specification is better suited for the data at hand.

[28] Vuong test strongly rejects the lognormal in favor of the truncated model in terms of fit with $p$ value of 0.000.

[29] For the dynamically complete panel data model (i.e., $D\left(M_{it} \mid \boldsymbol{X}_{it}, M_{it-1}, \boldsymbol{X}_{it-1}, \ldots\right) = D\left(M_{it} \mid \boldsymbol{X}_{it}\right)$ where $D\left(\cdot\right)$ denotes conditional distribution of $M_{it}$), inference is the same as would be for the independent cross sections.

Tab. 6.4: Pooled and Panel Bilateral FDI at Extensive and Intensive Margins

| Variables | LIC | | | MIC | | | Overall Sample | | |
|---|---|---|---|---|---|---|---|---|---|
| | Lognormal Hurdle | Exponential Type II Tobit | Mundlak Chamberlain Panel | Lognormal Hurdle | Exponential Type II Tobit | Mundlak Chamberlain Panel | Lognormal Hurdle | Exponential Type II Tobit | Mundlak Chamberlain Panel |
| **Intensive Margin of Log FDI** | | | | | | | | | |
| Log Destination Income | -1.41 | -1.872 | -1.153 | -1.54** | -1.143* | -1.284*** | -1.53*** | -1.230** | -1.182*** |
| | (1.208) | (1.262) | (-0.962) | (0.662) | (0.65) | (0.481) | (0.583) | (0.583) | (0.428) |
| Log Source Income | 2.37* | 1.965 | 3.608*** | 5.30*** | 4.673*** | 5.259*** | 5.00*** | 4.373*** | 4.989*** |
| | (1.418) | (1.484) | (1.104) | (0.797) | (0.885) | (0.577) | (0.701) | (0.776) | (0.511) |
| Log Lending Rate | -1.07** | -0.992* | -1.122*** | -0.07 | -0.17 | -0.230** | -0.156 | -0.243 | -0.297*** |
| | (0.424) | (0.542) | (0.342) | (0.136) | (0.187) | (0.0993) | (0.131) | (0.181) | (0.0954) |
| Log Labor productivity | 5.03*** | 6.582*** | 3.519*** | 2.44*** | 2.047** | 2.504*** | 2.650*** | 2.522*** | 2.509*** |
| | (1.651) | (2.038) | (1.341) | (0.745) | (0.805) | (0.548) | (0.678) | (0.773) | (0.503) |
| **Extensive Margin of FDI** | | | | | | | | | |
| Log Startup Costs | -0.189** | -0.185** | -0.188** | 0.0614 | 0.0586 | 0.0613 | -0.0287 | -0.0305 | -0.0281 |
| | (0.0854) | (0.0886) | (0.0854) | (0.0658) | (0.0657) | (0.0659) | (0.0493) | (0.0493) | (0.0493) |
| Log Lending Rate | 0.153 | 0.154 | 0.1506 | -0.0726 | -0.0722 | -0.0611 | -0.0596 | -0.0574 | -0.0465 |
| | (0.22) | (0.22) | (0.2205) | (0.101) | (0.101) | (0.101) | (0.0932) | (0.0931) | (0.093) |
| Log Labor productivity | 0.894 | 0.92 | 0.893 | 1.036*** | 1.041*** | 1.040*** | 0.882*** | 0.891*** | 0.888*** |
| | (0.551) | (0.566) | (0.55) | (0.272) | (0.273) | (0.273) | (0.238) | (0.238) | (0.238) |
| $\hat{\sigma}$ ($\hat{\sigma_u}$ / $\hat{\sigma_\epsilon}$) | 1.67*** | 1.70*** | 1.19/1.18 | 1.73*** | 1.73*** | 1.29***/1.17*** | 1.75*** | 1.75*** | 1.30***/1.17*** |
| | (0.0435) | (0.088) | (.081/.038) | (0.0216) | (0.025) | (.045/.017) | (.0197) | (0.0240) | (.040/.016) |
| $\hat{\rho}$ | - | 0.254 | .51 | - | 0.137 | 0.55 | - | 0.176 | 0.55 |
| | | | .040 | | | .019 | | | .017 |
| Observations | | | 687/2375 | | | 2962/5170 | | | 3649/7800 |
| Log likelihood | -4748.95 | | -4650.88 | -22304.34 | | -21720.51 | -27168.72 | | -26459.18 |

Note: time coverage is 2002-2010 The difference in observations in ET2T arises because all negative flows are truncated as zeroes, potentially carrying information on the existence of a link but absence of a positive capital flow. Mundlak-Chamberlain method includes time averages of relevant explanatory variables and inverse Mills ratios from the first stage Probit regressions. Time averages of startup costs and labor productivity are omitted in the first stage due to collinearity. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term and directional fixed effects (source and destination countries).

Tab. 6.5: Pooled and Panel Bilateral Portfolio at Extensive and Intensive Margins

| Variables | LIC | MIC | Overall Sample | |
|---|---|---|---|---|
| | Lognormal Hurdle | Lognormal Hurdle | Lognormal Hurdle | Exponential Type II Tobit |
| **Intensive Margin of Log Portfolio** | | | | |
| Log Destination Income | 0.290 | 0.0698*** | 0.0825*** | 0.0767** |
| | (0.235) | (0.0261) | (0.0259) | (0.0320) |
| Log Source Income | 0.127 | 1.067*** | 1.024*** | 0.930*** |
| | (0.338) | (0.0643) | (0.0634) | (0.0961) |
| Contract Risk | 0.425 | 0.463** | 0.481** | -0.243 |
| | (1.410) | (0.235) | (0.232) | (0.181) |
| **Extensive Margin of Portfolio** | | | | |
| Log Startup Costs | -0.117* | 0.329*** | 0.208*** | 0.289*** |
| | (0.0704) | (0.0480) | (0.0417) | (0.0498) |
| Contract Risk | -0.176 | 0.105 | 0.0807 | 0.195 |
| | (0.257) | (0.106) | (0.0987) | (0.298) |
| $\hat{\sigma}$ ($\hat{\sigma}_u$/$\hat{\sigma}_\varepsilon$) | 2.50e-17 /1.222*** | .624***/1.122*** | .6555***/1.13023*** | 1.294*** |
| | (.2818)/(.0713) | (.0359 )/(.0176) | (.0356 )/(.0174) | (.0276) |
| $\hat{\rho}$ | 4.19e-34 | .236 | .252 | .111 |
| | (9.44e-18) | (0.023) | (0.023) | |
| Observations | 147 | 2,591 | 2,738 | 4,932 |

Note: time coverage is 2002-2010. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term and directional fixed effects (source and destination countries).

yield a random effects tobit model. However, we are interested in the "dual" nature of capital flows and proceed to a two-step estimation procedure to unravel two margins. As before, a full sample probit estimation is followed by a selection equation carried out on the positive-flows subsample. The model assumes that different sets of variables could be used in the two-step estimations. Crucially, the second method, exponential type II tobit, allows for conditional correlation between selection and positive flow equations, even after controlling for observable covariates.

We further relax the crucial random effects assumption of independence of individual effects and independent variables and enable analysis when error terms in decision and volume equations are correlated. To achieve this in the panel probit model, we follow Papke and Wooldridge (2008) and use the Mundlak-Chamberlain approach. The unobserved effects are proxied by the observables which are time averages of the independent variables $\boldsymbol{X}_{it}$ and this helps dealing with more general unobserved heterogeneity.[30] However, we go further and consider a case when error terms in decision and volume equations are correlated. Using additional variables (time averages of explanatory variables) together with the $T$ inverse Mills ratios yields consistent estimates in the conditional FDI flow equation, provided the model is not misspecified. We report results of this procedure under the heading "Mundlak Chamberlain Panel".[31] Further robustness checks are given in Appendix D.

As is clear from Table 6.4, extensive margin (or the decision to invest) is affected by the startup costs (entry barriers) only for the LICs, consistently among estimation methods. Indeed, market entry barriers are consistent with many zeroes and absence of investment links. Further, labor productivity is a strong and consistent determinant of the decision to invest only for MICs and overall sample. Lending rate, though always negative, never affects investment decision (establishing a link) significantly. This can reflect the fact that financial markets are underdeveloped in low and middle income countries, therefore, making investors rely on access to financial markets elsewhere.

We further find that source income is a strong predictor of the volume of capital. Finally, labor productivity matters for the intensity, consistently across estimations.[32] This finding is supportive of supply side arguments - productivity at destination market, conditional on other factors, affects the intensity of direct investment.

Regarding portfolio flows, once both margins are considered, the effect of contract vanishes (see Table 6.5). However, it seems that portfolio flows act as substitutes for the FDI, since startup costs positively correlate with the probability to engage in FPI, after conditioning on other included variables. The puzzle of higher destination income conditionally attracting more portfolio investment persists. To draw more reliable inferences, we require more data on this type of investment, especially for the least developed countries. However, the potential interaction between FDI promises to be an important research direction which we hope to follow with new data arrivals.

In summary, zeroes of investment flows require both theoretical and empirical treatment. We observe

---

[30] Moreover, this approach is somewhat similar to the recent contribution by Pesaran (2006) where cross sectional averages are used to mimic common factors, similarly to global approach in Dees, Mauro, Pesaran, and Smith (2007), where global unobserved factor can be approximated by observable cross sectional averages.

[31] An alternative to this approach is that of Kyriazidou (1997). The method relies on differencing out the selection and unobservable individual effects. The proposed estimator is of nonparametric nature and requires kernel weights. The choice of kernel density function and a sequence of bandwidths is not trivial. Moreover, the window widths crucially affect the variance of the limiting distribution of the estimator and sensitivity is an issue for smaller samples. Therefore, we resort to a parametric approach which allows for sufficiently general unobserved heterogeneity.

[32] An interesting observation - when we include marginal product of capital (see Appendix) - the correlation between the decision to invest and the volume choice varies significantly among country groups. A strong negative correlation exists for LICs, stating that the unobserved factors which affect the investment decision have a strong negative effect on how much to invest. There is no statistical relationship for the MICs; a moderate positive correlation between decisions exists for overall sample. We apply both pooled probit which is consistent estimator with the cluster-robust errors to correct for error correlation over time. Alternatively, one can fit the random effects model proposed by Butler and Moffitt (1982) which still maintains the homoskedasticity assumption but extends the pooled model by allowing cross period correlation. However, the gain in panel setting is marginal because random effects maintain the assumption of independence in individual effects and explanatory variables.

that ignoring extensive margin affects estimates and their statistical significance. Most importantly, modeling two margins reveals differences among developing countries. This is critical for both, understanding the current developments and drawing policy implications, in particular across development stages.

## 7    Robustness

We deviate from the baseline model, and consider a number of robustness checks. We analyze how results change once investment threshold is allowed to be non-zero. Further, despite accounting for the two margins, we investigate more flexible dependence structures than are admitted by our previous techniques. Finally, we also check other theory implied results for both FDI and portfolio flows which at least partly confirm the validity of theoretical specification. Further, in Appendix we experiment with pseudo Poisson maximum likelihood as proposed by Santos-Silva and Tenreyro (2007) and marginal product instead of lending rate.

### 7.1    Stability and Endogenous Investment Threshold

One of the main problems regarding statistical treatment of FDI lies in its very definition that requires an equity stake of 10% or more for an investment to be classified as FDI. In general, FDI itself has three components: equity capital, intra-firm loans, and reinvestment of retained earnings. Because of this definition, it is not clear if all zeroes are indeed unique - there might be "marginal" investments which are not reported. Moreover, there might be measurement errors that produce this result instead of being a deliberate choice of investors, which would be at odds with our theoretical motivation.

   We start by considering how parameters behave once we model different quantiles of capital flows, rather than its conditional mean. This helps to avoid parametric assumptions about the error term. We plot in Figure 7.1 and Figure 7.2 the bootstrapped parameters and their confidence intervals from both FDI and portfolio quantile (median) regressions (without and with directional fixed effects to otherwise baseline specifications). For FDI flows, the effects exerted by source are always positive, mainly positive for labor productivity (in particular, for LICs),[33] and varies for lending rate depending on the quantile (tends to be negative for LICs). Destination income has mainly negative effect only if directional fixed effects are accounted for (thus capital intensity increases with lower development, but only conditional on fixed effects which mimic multilateral resistance terms). Intuition is mainly confirmed - the investor should be more developed to engage in FDI whereas more opportunities are found at destinations which have a conditionally higher labor productivity and conditionally lower income. Portfolio flows are positively related to source income but very weakly related to destination income and contract risk, both effects being largely dependent on the quantile. Again, failure to account for the directional fixed effects would have led to erroneous conclusions of source income exerting no (or slightly negative) effects on portfolio flows.

   Another problem related to parameter stability concerns the nature of zero flows. Eaton and Tamura (1994) developed an early solution to incorporate zeros that can be thought of as a model of $\ln(M_{ni} + a)$ where instead of arbitrarily setting $a = 1$, it is instead treated as a parameter to be estimated. In our context, one could think of $a$ as a minimum investment level that is needed to be incurred before the flow is accounted for. Effectively, this is an extended tobit which defines a strictly positive latent variable $M_{ni}^{\star}$ and a threshold $\hat{a}$ which, unfortunately, lacks a compelling structural interpretation. Another drawback of ET Tobit is that it is not a "canned" program. Tobit's consistency requires homoskedastic errors, an assumption barely met in a multi-country setting with countries of different sizes, development levels, and other characteristics, both observable and unobservable.[34] Martin and Pham (2008) consider

---

[33] We also graph parameters for LIC and MIC countries separately, however, report results for a full sample. Other graphs are available from authors.

[34] Eaton and Kortum (2001) propose another method that has the advantage of being both easier to implement and interpret. Suppose that there is minimum level of investment, $M_{ni}^{\min}$, such that if "ideal" capital flow, $M_{ni}^{\star}$, falls below $M_{ni}^{\min}$

Fig. 7.1: Simple OLS (horizontal lines) and quantile coefficients and their confidence intervals for FDI (left) and portfolio (right) flows without directional fixed effects
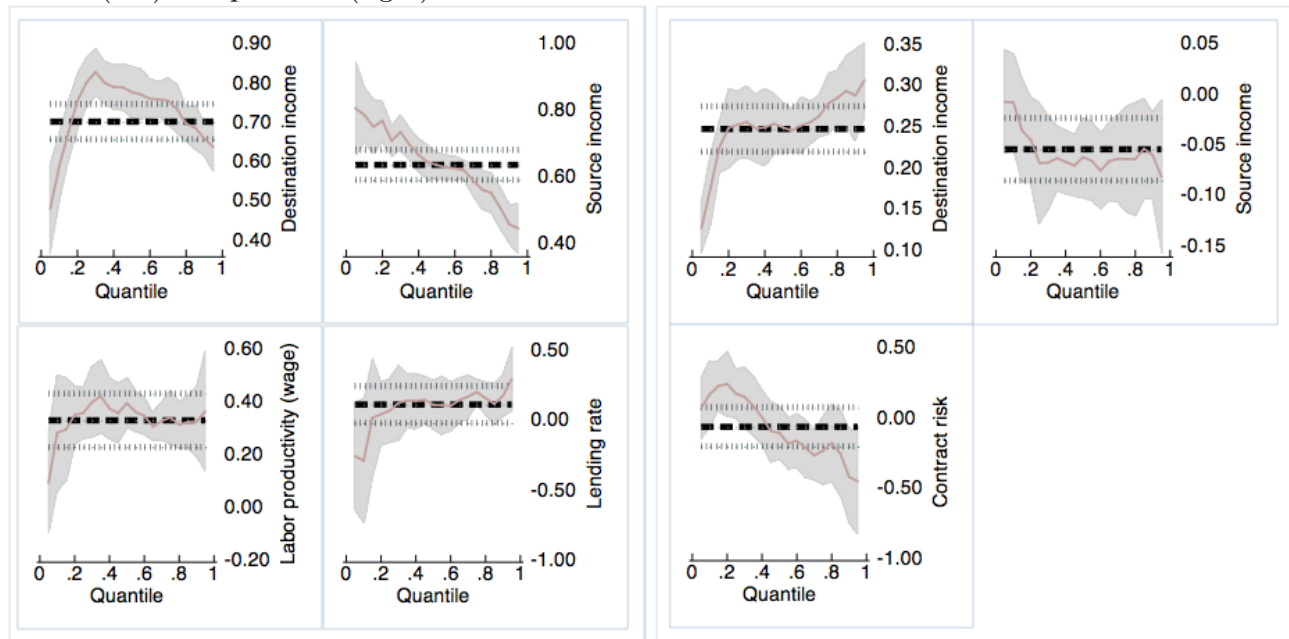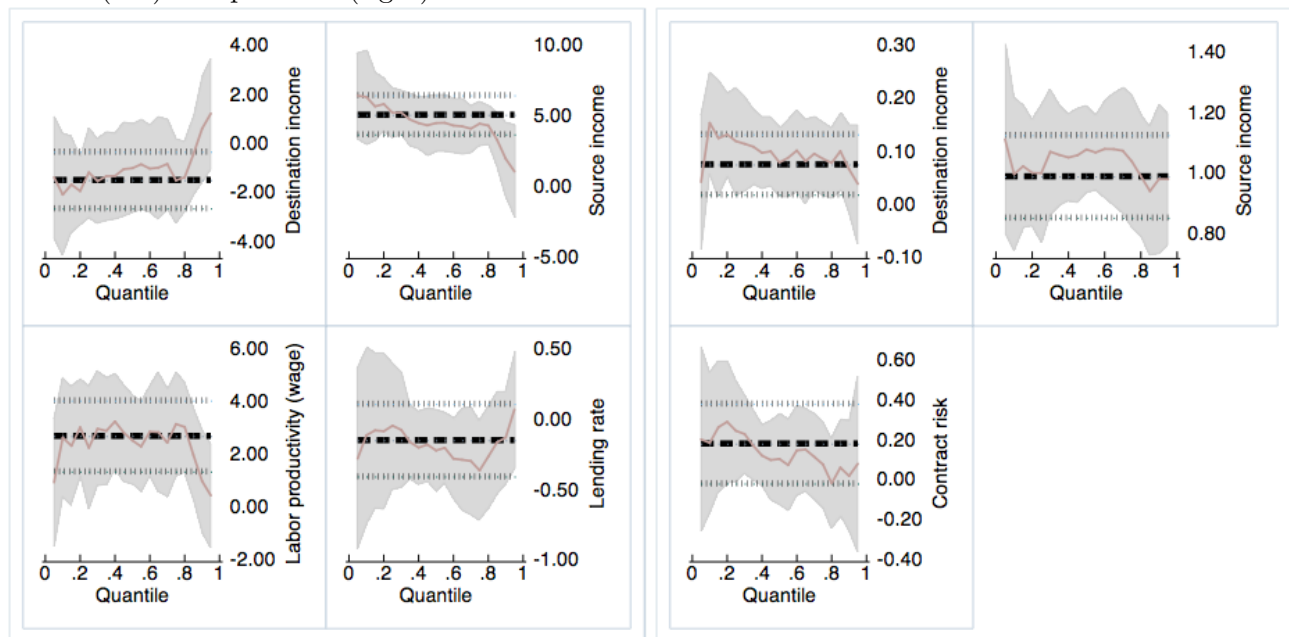


Fig. 7.2: Simple OLS (horizontal lines) and quantile coefficients and their confidence intervals for FDI (left) and portfolio (right) flows with directional fixed effects

DGPs involving threshold values and find that tobit and Heckman methods outperform the Poisson PML (though see Appendix B.2.1 for Poisson regressions). The problem with Tobit is that if the error term is heteroskedastic, such that bilateral trade has a constant variance to mean ratio, it is inconsistent whereas Poisson or multinomial PMLs are unbiased.

Kourtellos, Stengos, and Tan (2009) propose estimating a threshold parameter based on a concentrated least squares method that involves an inverse Mills ratio bias correction term in each regime. The method is coined as Structural Threshold Regression and is composed of two regressions for two regimes. We allow for relationship between the variables of interest in each of the two regimes and use different threshold variables. To operationalize this approach, we implement successive Chow tests for endogenous thresholds when significant relationship starts appearing. Our results indicate that the behavior of the main equations remain stable across different levels of capital flows and development stages of receiving economies.[35] Therefore, our results should not suffer from the misspecification of investment threshold in two-tier (intensive and extensive) estimations. However, parametric assumptions still need to be addressed to which we turn next.

## 7.2  Copulae-based Heckman Method

As a robustness check, we resort to the recent contribution by Hasebe and Vijverberg (2012). Parametric approach is criticized for its sensitivity to the distributional assumption. The alternative involves a copula where a multivariate distribution is constructed from separately specified marginal distributions. The Generalized Tukey Lambda (GTL)-Copula approach inserts more flexible marginal distribution into the copula function. The GTL distribution is a versatile univariate distribution that permits a wide range of skewness of thick- or thin-tailed behavior in the data that it represents and, therefore, is a good candidate distribution for modeling the unobservables in the extensive and intensive margin equations. GTL-copula distribution effectively frees the extensive margin's model from any particular distributional assumption.

What is more, this flexibility is achieved with just a few additional parameters, which is both parsimonious and time-efficient relative to semi- or non-parametric approaches. The collective set of copulas accommodates diverse dependence structures between two random variables. Unlike the traditional estimator (the normal-Gaussian estimator), the GTL-copula estimator is much less dependent on the presence of an instrument in the selection equation that fulfills the exclusion restriction, no longer making it problematic that the extensive margin equation contains the same explanatory variables as the capital volume equation.

We experiment with different copula functions governing the dependence between the errors in the capital flows equation and extensive margin. In Table 7.1, we report Gaussian copula results with probit (Normal) marginal distribution of the error term in the extensive margin. Though we do not report, in addition to Normal, we also allow for errors to have Student's t marginal distributions in the capital flows equation. Our main predictions are in line with previous results, the capital tends to flow to countries with larger income, after conditioning on other theory implied variables.

Table 7.2 collects information on portfolio flows. It confirms the importance of startup costs and contract risk for establishing a portfolio investment link for the low income countries only. The startup costs preserve the significance across development stages unlike the contract risk. Higher destination income creates more opportunities and therefore makes investment larger in magnitude whereas source income, unlike FDI, plays no important role (if anything, less prosperous investors seem to favour portfolio investment).

---

we observe $M_{ni} = 0$ but otherwise we observe $M_{ni} = M_{ni}^{\star}$. To estimate the model, all the observed zeros are replaced with minimum value $M_{ni}^{\min}$ for each bilateral link and the new bottom-coded $\ln M_{ni}$ is the dependent variable in a Tobit command that allows for a user-specified lower limit of $\ln M_{ni}^{\min}$. Method's advantage is easy implementation and no requirement of exclusion restrictions.

[35] Results are not reported but available upon request.

Tab. 7.1: Copula-based FDI Flows Model

|  | LIC | MIC | Overall Sample |
|---|---|---|---|
| Variables | | Copula | |
| **Intensive Margin of Log FDI** | | | |
| Log Destination Income | 0.448*** | 0.681*** | 0.683*** |
|  | (0.143) | (0.0517) | (0.0407) |
| Log Source Income | 0.406*** | 0.599*** | 0.561*** |
|  | (0.0961) | (0.0453) | (0.0419) |
| Log Lending Rate | 0.254 | 0.502*** | 0.589*** |
|  | (0.307) | (0.163) | (0.149) |
| Log Labor productivity | 0.0870 | 0.720*** | -0.280** |
|  | (0.245) | (0.221) | (0.117) |
| **Extensive Margin of FDI** | | | |
| Log Startup Costs | -0.0693** | -0.122*** | -0.0985*** |
|  | (0.0347) | (0.0223) | (0.0185) |
| Log Lending Rate | -0.267*** | -0.0760 | -0.153*** |
|  | (0.0895) | (0.0526) | (0.0458) |
| Log Labor productivity | -0.0566 | -0.256*** | 0.134*** |
|  | (0.0675) | (0.0603) | (0.0329) |
| Log $\sigma$ | 1.2456*** | 1.2382*** | 1.2730*** |
|  | (0.1091) | (0.0390) | (0.0373) |
| Ancillary $\theta$ | -1.3335*** | -1.6772*** | -1.6252*** |
|  | (0.2146) | (0.0917) | (0.0864) |

Note: time coverage is 2002-2010. $\theta$ refers to the dependence parameter in the copula framework. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

Tab. 7.2: Copula-based Portfolio Flows Model

|  | LIC | MIC | Overall Sample |
|---|---|---|---|
| Variables | | Copula | |
| **Intensive Margin of Log Portfolio** | | | |
| Log Destination Income | 0.1435** | 0.2178*** | 0.2331*** |
|  | (0.0694) | (0.0273) | (0.0259) |
| Log Source Income | 0.0308 | -0.0628* | -0.0614* |
|  | (0.0628) | (0.0345) | (0.0332) |
| **Extensive Margin of Portfolio** | | | |
| Contract Risk | -0.3791*** | -0.0070 | 0.0821** |
|  | (0.0490) | (0.0545) | (0.0407) |
| Log Startup Costs | -0.0901*** | -0.1278*** | -0.2865*** |
|  | (0.0257) | (0.0246) | (0.0179) |
| Log $\sigma$ | 0.6430*** | 0.8477*** | 0.8103*** |
|  | (0.1489) | (0.1165) | (0.0429) |
| Ancillary $\theta$ | -0.7164*** | -0.4985 | -0.3734** |
|  | (0.2672) | (0.3630) | (0.1452) |

Note: time coverage is 2002-2010. $\theta$ refers to the dependence parameter in the copula framework. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

Tab. 7.3: Theory Implied Relationships: Ratio and Sum of Capital Flows

| Variables | Log Ratio ($\frac{M_{ni}^I}{M_{ni}^{FDI}}$) | | Log Sum ($M_{ni}^I + M_{ni}^{FDI}$) | |
| | FE (with incomes) | FE (without incomes) | FE (with incomes) | FE (without incomes) |
|---|---|---|---|---|
| Log Destination | -0.4323 | - | -1.7425*** | - |
| Income | (0.6026) | - | (0.4870) | - |
| Log Source | 1.0240 | - | 4.9515*** | - |
| Income | (1.2296) | - | (0.8850) | - |
| Log Lending | 0.1717 | 0.1529 | -0.2824 | -0.5286 |
| Rate | (0.9028) | (0.8982) | (0.2888) | (0.3528) |
| Log Labor | -1.9304** | -1.7727*** | 4.0875*** | 5.1349*** |
| Productivity | (0.8566) | (0.6504) | (0.5759) | (0.4440) |
| Log Startup | 0.6402** | 0.6049** | 0.0022 | -0.1737 |
| Costs | (0.2514) | (0.2401) | (0.0915) | (0.1302) |
| Contract | -0.0518 | -0.0631 | 0.0778 | -0.0249 |
| Risk | (0.1502) | (0.1544) | (0.0979) | (0.1236) |
| Observations | 1301 | 1301 | 1855 | 1855 |
| Within $R^2$ | 0.0156 | 0.0150 | 0.1622 | 0.1313 |
| $F$ test (incomes irrelevant) | $F(2, 370) = 0.35, p = 0.7062$ | | $F(2, 423) = 15.84, p = 0.0000$ | |

Note: time coverage is 2002-2010. FE stands for fixed effects with the reported Driscoll-Kraay standard errors which help controlling for spatial (cross-sectional) dependence. Inverse Mill's ratios are included. Standard errors in parentheses; *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

## 7.3 Theory Implied Relationship Between FDI and FPI

Recall the sum and ratio of the two types of investment, as reported in equations (5.2) and (5.3). The sum implies that the main driving forces are source and destination income, with all other terms entering as one nonlinear term whereas the ratio makes sure that the effects of income are eliminated.

To have a glimpse about this prediction, we construct a new variable which is a ratio of bilateral capital flows (portfolio to FDI). We take logs of (5.2) and (5.3) and consider random and fixed effects regressions for both specifications. Using robust (applicable with cluster-robust standard errors) Hausman test, as outlined in Wooldridge (2010), we conclude that fixed effects are preferred.[36]

Table 7.3 describes results once fixed effects and Mill's ratios are controlled for. The income effects disappear in the ratio and prevails in the sum, as predicted by theory. Labor productivity dictates the capital flow - it increases the sum and decreases the portfolio investment. Hence, FDI is attracted, conditional on other regressors, to countries with higher labor productivity. Startup costs make portfolio investment more attractive. There is some evidence that more conditional opportunities arise in less developed countries. Neither lending rate nor contract risk seem to play a significant role when both capital types are considered together.

## 8 Conclusions

Today's economic landscape is quickly changing and calls for an enquiry into global adjustments and capital flows, in particular with regard to low–income countries (LICs). To address recent intensification of capital flows to LICs, we analyze intensive and extensive margins of capital flows (the size of a flow versus a number of investors and receiving countries). Our motivation involved many zeroes in capital flows, in particular in early 2000s. The network structure has considerably changed along both dimensions. We set a theoretical model where the investment decision involves a share of costs which is fixed in

---

[36] The null of no systematic differences is rejected in both cases. In ratio case, $F(4, 475) = 7.00$ whereas in the sum case $F(6, 581) = 41.48$. The associated $p$ values are zero.

nature thereby generating two-stage decision and path dependency. FDI is portrayed as a supply whereas portfolio as a demand stories. We then map the theoretical account into an econometric framework, allowing for flexible estimation methods. Ignoring extensive margin affects estimates and their statistical significance. Results indicate that market entry costs do statistically affect decision to invest only for the LICs, consistently among estimation methods. This result, however, depends on whether dynamics is allowed for. Once lagged value of extensive margin is introduced, the effect of startup costs is dampened if not lost. Generally, market entry barriers are consistent with many zeroes and absence of links, and most probably reflect institutional environment, as emphasized in Alfaro, Kalemli-Ozcan, and Volosovych (2008).

LICs do not follow neoclassical arguments in neither of specifications and call for further investigation. Lending rate, though always negative, never affects investment decision (establishing a link) significantly. Labor productivity turns out to be a strong and consistent determinant of the decision to invest only for MICs and overall sample but matters for the intensity, consistently across estimations. At intensive margin, source income is a strong predictor of the volume of capital. Capital costs (lending rate) negatively affects the volume of capital, though is less robust predictor. Labor productivity matters for the intensity, consistently across estimations.

These results, demonstrating differences at two investment margins for different country groups, call for more research along the proposed lines. We foresee an extension with an explicit role for bilateral trade to be particularly promising.[37] This should be particularly easy as trade flows can be modeled using the same gravity framework (despite details of microfoundations, under mild qualifications, trade yields gravity relationship, see Head and Mayer, 2015).

Seeking a parsimonious approach, we have abstracted from a number of important issues. One relates to the risks of volatility and its effect on growth prospects (see contribution by Aghion and Banerjee (2005) whose framework might be useful to introduce capital flows). Also, abundance of natural resources constitutes an important source of attractiveness to invest and need a more complex modeling environment where, as in traditional Hecksher-Ohlin economy, endowment plays a role in determining a capital flow's direction and intensity.[38] Statistically, we are still constrained in time dimension and foresee many more explorations with explicit role for dynamics, especially from network's perspective, once more and better quality data become available. In such a case, we would be able to track transmission of shocks and spatio-temporal evolution capital movements, drawing from the current contribution due to Rebucci, Cesa-Bianchi, Pesaran, and Xu (2012) which analyze how changes in trade linkages between China, Latin America, and the rest of the world have altered the transmission mechanism of international business cycles to Latin America.[39] This is an important extension needed for tracking macro-financial stability and transmission of macro policies globally, in particular to more fragile low income countries.

---

[37] One advantage from macroeconomic perspective is the analysis of Balance of payments, BoP. We have analyzed capital account but still lack account and interaction of another, current account. Recall that BoP = Current Account + Capital Account + Error Term (balancing item).

[38] Esfahani, Mohaddes, and Pesaran (2012, 2014) draw long run implications of a resource discovery, unlike the standard literature on the "Dutch disease" and the "resource curse". Real output in the long run is shaped by oil exports through their impact on capital accumulation, and the foreign output as the main channel of technological transfer.

[39] The evidence reported also suggests that the emergence of China as an important source of world growth might be the driver of the so called "decoupling" of emerging markets business cycle from that of advanced economies reported in the existing literature.

## References

AGHION, P., AND A. BANERJEE (2005): *Volatility and Growth*, Clarendon Lectures in Economics. OUP Oxford.

ALFARO, L., S. KALEMLI-OZCAN, AND V. VOLOSOVYCH (2008): "Why Doesn't Capital Flow from Rich to Poor Countries? An Empirical Investigation," *The Review of Economics and Statistics*, 90(2), 347–368.

ANDERSON, J. E., AND E. VAN WINCOOP (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93(1), 170–192.

AOYAMA, H., Y. FUJIWARA, AND Y. IKEDA (2010): *Econophysics and companies: statistical life and death in complex business networks*. Cambridge University Press.

ARVIS, J.-F., AND B. SHEPHERD (2011): "The Poisson quasi-maximum likelihood estimator: A solution to the "adding up" problem in gravity models," Discussion paper, MPRA Paper No. 34334.

ATALAY, E., A. HORTACSU, J. ROBERTS, AND C. SYVERSON (2011): "Network structure of production," *Proceedings of the National Academy of Sciences*.

AXTELL, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293(5536), 1818–1820.

BAIER, S. L., AND J. H. BERGSTRAND (2009): "Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation," *Journal of International Economics*, 77(1), 77–85.

BEHAR, A., AND B. D. NELSON (2009): "Trade Flows, Multilateral Resistance and Firm Heterogeneity," Discussion paper, University of Oxford.

BERGSTRAND, J. H., P. EGGER, AND M. LARCH (2008): "Gravity Redux: Structural Estimation of Gravity Equations with Asymmetric Bilateral Trade Costs," .

BERNARD, A. B., S. J. REDDING, AND P. K. SCHOTT (2007): "Comparative advantage and heterogeneous firms," *The Review of Economic Studies*, 74(1), 31–66.

BURNSIDE, C., AND A. TABOVA (2009): "Risk, Volatility, and the Global Cross-Section of Growth Rates," NBER Working Papers 15225, National Bureau of Economic Research, Inc.

BUTLER, J. S., AND R. MOFFITT (1982): "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model," *Econometrica*, 50(3), 761–64.

CHANEY, T. (2008): "Distorted Gravity: the Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–1721.

COEURDACIER, N., AND P. MARTIN (2009): "The geography of asset trade and the euro: Insiders and outsiders," *Journal of the Japanese and International Economies*, 23(2), 90 – 113, Special Conference Issue on Financial Globalization (20th Anniversary Trio Conference).

CORTE, P. D., L. SARNO, AND G. SESTIERI (2012): "The Predictive Information Content of External Imbalances for Exchange Rate Returns: How Much Is It Worth?," *The Review of Economics and Statistics*, 94(1), 100–115.

CRAGG, J. G. (1971): "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods," *Econometrica*, 39(5), 829–844.

CROZET, M., AND P. KOENIG (2010): "Structural gravity equations with intensive and extensive margins," *Canadian Journal of Economics/Revue canadienne d'économique*, 43(1), 41–62.

DEES, S., F. D. MAURO, M. H. PESARAN, AND L. V. SMITH (2007): "Exploring The International Linkages of the Euro Area: A Global VAR Analysis," *Journal of Applied Econometrics*, 22(1), 1–38.

DUAN, N. (1983): "Smearing Estimate: A Nonparametric Retransformation Method," *Journal of the American Statistical Association*, 78(383), pp. 605–610.

EATON, J., AND S. KORTUM (2001): "Trade in capital goods," *European Economic Review*, 45(7), 1195–1235.

EATON, J., AND A. TAMURA (1994): "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns," *Journal of the Japanese and International Economies*, 8(4), 478–510.

EGGER, P., AND D. M. RADULESCU (2011): "Labor Taxation and Foreign Direct Investment," *Scandinavian Journal of Economics*, 113(3), 603–636.

EICKER, F. (1963): "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions," *The Annals of Mathematical Statistics*, 34(2), pp. 447–456.

ESFAHANI, H. S., K. MOHADDES, AND M. H. PESARAN (2012): "Oil Exports and the Iranian Economy," Discussion paper, University of Cambridge.

———— (2014): "An Empirical Growth Model for Major Oil Exporters," *Journal of Applied Econometrics*, 29(1), 1–21.

FEENSTRA, R. (2004): *Advanced International Trade: Theory and Evidence*. Princeton University Press.

FELBERMAYR, G. J., AND W. KOHLER (2006): "Exploring the Intensive and Extensive Margins of World Trade," *Review of World Economics*, 142, 642–674, 10.1007/s10290-006-0087-3.

GANDOLFO, G. (2002): *International Finance and Open-Economy Macroeconomics*. Springer.

GOURINCHAS, P.-O., AND H. REY (2007): "International Financial Adjustment," *Journal of Political Economy*, 115(4), 665–703.

GREENE, W. H. (2003): *Econometric Analysis*. New York: Pearson Education.

HASEBE, T., AND W. P. VIJVERBERG (2012): "A Flexible Sample Selection Model: A GTL-Copula Approach," IZA Discussion Papers 7003, Institute for the Study of Labor (IZA).

HEAD, K., AND T. MAYER (2015): "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by K. R. Elhanan Helpman, and G. Gopinath, vol. 4 of *Handbook of International Economics*, pp. 131 – 195. Elsevier.

HECKMAN, J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–161.

HELPMAN, E., J. M. MELITZ, AND S. R. YEAPLE (2004): "Export Versus FDI with Heterogeneous Firms," *American Economic Review*, 94(1), 300–316.

HELPMAN, E., M. J. MELITZ, AND Y. RUBINSTEIN (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123(2), 441–487, 2007.

IMF (2011): *Regional Economic Outlook, April 2011: Sub-Saharan Africa - Recovery and New Risks*, World Economic and Financial Surveys Series. International Monetary Fund.

JIA, S., AND S. RATHI (2008): "On Predicting Log-transformed Linear Models with Heteroscedasticity," *SAS Global Forum*.

JONES, A. M. (1992): "A Note on Computation of the Double-Hurdle Model with Dependence with an Application to Tobacco Expenditure," *Bulletin of Economic Research*, 44(1), 67–74.

KLEINERT, J., AND F. TOUBAL (2010): "Gravity for FDI," *Review of International Economics*, 18(1), 1–13.

KOCH, W., AND J. P. LESAGE (2009): "Latent Multilateral Trade Resistance Indices: Theory and Evidence," Discussion paper.

KOURTELLOS, A., T. STENGOS, AND C. M. TAN (2009): "Structural Threshold Regression," Working Paper Series 22, The Rimini Centre for Economic Analysis.

KYRIAZIDOU, E. (1997): "Estimation of a Panel Data Sample Selection Model," *Econometrica*, 65(6), pp. 1335–1364.

LASTAUSKAS, P., AND M. H. PESARAN (2013): "Openness and Technological Interdependence in Global Economy," Discussion paper, University of Cambridge.

LEVY YEYATI, E., E. STEIN, AND C. DAUDE (2003): "Regional Integration and the Location of FDI," Working Papers 492, Inter-American Development Bank.

LOWE, M., C. PAPAGEORGIOU, AND F. PEREZ-SEBASTIAN (2012): "The Public and Private MPK," DEGIT Conference Papers c017(021), DEGIT, Dynamics, Economic Growth, and International Trade.

MARTIN, P., AND H. REY (2004): "Financial super-markets: size matters for asset trade," *Journal of International Economics*, 64(2), 335–361.

MARTIN, W., AND C. S. PHAM (2008): "Estimating the Gravity Model When Zero Trade Flows are Frequent," Economics Series 3, Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

MERTON, R. C. (1969): "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *The Review of Economics and Statistics*, 51(3), pp. 247–257.

MIRANDA, A. (2007): "Dynamic probit models for panel data: A comparison of three methods of estimation," United Kingdom Stata Users' Group Meetings 2007 11, Stata Users Group.

MODY, A. (2007): *Foreign Direct Investment And the World Economy*, Routledge Studies in the Modern World Economy. Routledge.

MODY, A., AND A. P. MURSHID (2005): "Growing up with capital flows," *Journal of International Economics*, 65(1), 249–266.

NEWEY, W. K. (1985): "Maximum Likelihood Specification Testing and Conditional Moment Tests," *Econometrica*, 53(5), 1047–70.

OBSTFELD, M., AND K. ROGOFF (2000): *The Six Major Puzzles in International Macroeconomics. Is There a Common Cause?* pp. 339–390. MIT Press.

OKAWA, Y., AND E. VAN WINCOOP (2012): "Gravity in International Finance," *Journal of International Economics*, 87(2), 205–215.

PAPKE, L. E., AND J. M. WOOLDRIDGE (2008): "Panel data methods for fractional response variables with an application to test pass rates," *Journal of Econometrics*, 145(1-2), 121–133.

PESARAN, M. H. (2006): "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74(4), 967–1012.

RAMSEY, J. B. (1969): "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis," *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(2), pp. 350–371.

RAZIN, A., Y. RUBINSTEIN, AND E. SADKA (2005): "Corporate Taxation and Bilateral FDI with Threshold Barriers," NBER Working Papers 11196, National Bureau of Economic Research, Inc.

RAZIN, A., AND E. SADKA (2007a): *Foreign Direct Investment: Analysis of Aggregate Flows*. Princeton University Press.

——— (2007b): "Introduction to Foreign Direct Investment: Analysis of Aggregate Flows," in *Foreign Direct Investment: Analysis of Aggregate Flows*, Introductory Chapters. Princeton University Press.

——— (2007c): "Productivity and Taxes as Drivers of FDI," Working Papers 172007, Hong Kong Institute for Monetary Research.

RAZIN, A., E. SADKA, AND H. TONG (2008): "Bilateral FDI Flows: Threshold Barriers and Productivity Shocks," *CESifo Economic Studies*, 54(3), 451–470.

REBUCCI, A., A. CESA-BIANCHI, M. H. PESARAN, AND T. XU (2012): "China's Emergence in the World Economy and Business Cycles in Latin America," *Journal of LACEA Economia.*

ROBINSON, P. M. (1982): "On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables," *Econometrica*, 50(1), 27–41.

ROY, J. R. (2004): *Spatial Interaction Modeling: A Regional Science Context.* Berlin: Springer-Verlag.

SAMUELSON, P. A. (1969): "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *The Review of Economics and Statistics*, 51(3), 239–46.

SANTOS-SILVA, J., AND S. TENREYRO (2007): "The log of gravity," *Review of Economics and Statistics*, 88, 641–658.

SILVA, J. M. C. S., AND S. TENREYRO (2009): "Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically," CEP Discussion Papers dp0935, Centre for Economic Performance, LSE.

STEWART, M. (2006): "Maximum simulated likelihood estimation of random-effects dynamic probit models with autocorrelated errors," *Stata Journal*, 6(2), 256–272.

STEWART, M. B. (2007): "The interrelated dynamics of unemployment and low-wage employment," *Journal of Applied Econometrics*, 22(3), 511–531.

TABOVA, A. (2010): "Portfolio Diversification and the Cross-Sectional Distribution of Foreign Investment," Discussion paper, Duke University.

TEEKEN, R., AND J. KOERTS (1972): "Some Statistical Implications of the Log Transformation of Multiplicative Models," *Econometrica*, 40(5).

TILLE, C., AND E. VAN WINCOOP (2010): "International capital flows," *Journal of International Economics*, 80(2), 157–175.

WELCH, I. (1992): "Sequential Sales, Learning, and Cascades," *The Journal of Finance*, 47(2), pp. 695–732.

WHITE, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48(4), 817–38.

WOOLDRIDGE, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, second edn.

YOTOV, Y. V., AND M. OLIVERO (2012): "Dynamic Gravity: Endogenous Country Size and Asset Accumulation," *Canadian Journal of Economics*, 45(1).

# Appendix

## A   Derivations

### A.1   Firms' Decisions

The firms take the residual demand in (4.2) as given and maximize profits by choosing a price. In any country $i$, the profit a $\varphi-$firm is given by

$$\max_{\{p_i(\omega)\}\geq 0} \pi_i(\varphi) = (p_i(\omega) - v_i \tau_{in} \varphi) \left( Q_i \left( \frac{p_i(\omega)}{P_i} \right)^{-\sigma} \right) - v_i f_{in}, \tag{A.1}$$

where the first order condition leads to

$$\left( 1 - \sigma + \sigma v_i \tau_{in} \varphi p_i(\omega)^{-1} \right) \left( Q_i \left( \frac{p_i(\omega)}{P_i} \right)^{-\sigma} \right) = 0,$$

and produces the price $p_{in}(\omega) = \frac{\sigma}{\sigma-1} v_i \tau_{in} \varphi$. Notice that the revenue is given by

$$R_{in}(\varphi) \equiv p_{in}(\omega) q_{in}(\omega) = p_{in}(\omega)^{1-\sigma} Q_n P_n{}^\sigma$$
$$= \left( \frac{\sigma}{\sigma-1} c_{in}(\varphi) \right)^{1-\sigma} Q_n P_n{}^\sigma = \left( \frac{\sigma}{\sigma-1} c_{in}(\varphi) \right)^{1-\sigma} \frac{P_n Q_n}{P_n{}^{1-\sigma}}.$$

Plugging price into profit yields $\pi_{in}(\varphi) = \frac{p_{in}(\varphi)q_{in}(\varphi)}{\sigma} - v_i f_{in} = \frac{R_{in}(\varphi)}{\sigma} - v_i f_{in}$. Finally, the payment to the factors of production are uncovered from the cost minimization problem,

$$\min_{\{l_i(\varphi)\}\geq 0, \{k_i(\varphi)\}\geq 0} r_i k(\varphi) + w_i l(\varphi) = C_{in}(q) = \frac{r_i{}^{\alpha_i} w_i{}^{1-\alpha_i}}{\alpha_i{}^{\alpha_i}(1-\alpha_i)^{1-\alpha_i}} (q\tau_{in}\varphi + f_{in})$$
$$\text{s.t. } y = \frac{k^{\alpha_i} l^{1-\alpha_i}}{\varphi} = q_i(\omega) = Q_i \left( \frac{p_i(\omega')}{P_i} \right)^{-\sigma}.$$

The associated Lagrangian yields the results as in the main text. However, a simpler approach is to observe that with the competitive factor markets, the payments are just marginal costs

$$w_i l_{in}(\varphi) = (1-\alpha_i) \frac{\sigma-1}{\sigma} p_{in}(\varphi) y_{in}(\varphi) = (1-\alpha_i) \frac{\sigma-1}{\sigma} R_{in}(\varphi),$$
$$r_i k_{in}(\varphi) = \alpha_i \frac{\sigma-1}{\sigma} p_{in}(\varphi) y_{in}(\varphi) = \alpha_i \frac{\sigma-1}{\sigma} R_{in}(\varphi).$$

### A.2   Agents' Decisions

We build on Burnside and Tabova (2009); Tabova (2010) and Okawa and van Wincoop (2012) and demonstrate how the FDI model can be extended to portfolio flows in a global economy with $\mathbb{N}$ countries. Following the latter contribution, we define $\mathbb{N}+2$ assets where the first $\mathbb{N}$ assets are country-specific risky assets, also a risk-free bond that is in zero net supply and an asset whose return is perfectly correlated with the global shock which is in zero net supply. A simplifying assumption is the existence of the global asset which allows hedging against the global risk factor, so that the only risk that matters for portfolio allocation across the $\mathbb{N}$ equity is the country-specific risk. Let the amount to be invested be $W_i$, total income $E_i$, consumption $Q_i$, the price of an asset $P_i^I$ and $I_i$ is the country $i$'s equity. A simple accounting relation is described by

$$W_i = E_i - P_i Q_i = P_i^I I_i.$$

The decision space includes consumption, allocation of wealth across $\mathbb{N}+2$ assets, and production. The agent's budget constraint is

$$W_i R_i = (E_i - P_i Q_i) R_i,$$

where $R_i$ is the mean return in country $i$'s assets. Then, the portfolio return is a weighted average,

$$R_i^p = \sum_{n=1}^{\mathbb{N}} w_{ni} R_j + w_{ig} R_g + w_{if} R_f,$$

where $w_{in}$ describes country $i$'s fraction invested in country $n$, $w_{ig}$ is the fraction invested in the global asset and $w_{if}$ is the share invested in the risk-free asset. It is assumed that due to differences in language and regulatory systems, and easier access to local information, domestic agents are more informed than foreigners about the idiosyncratic payoff innovations on domestic equity claims. From the perspective of agents in country $i$, $\epsilon_n$ has a mean of 0 and variance

$$\tau_{ni} \sigma_n^2,$$

where information asymmetry is captured by $\tau_{ni} > \tau_{nn}$ when $n \neq i$.

### A.2.1  Optimal Agents' Decisions

Denote $w_{ii}$ the share invested in the domestic economy, $w_{ni}$ the vector of shares invested in $i$ from $j$.[40] Then the fraction employed in the riskless activity, $w_{if}$ , is:

$$w_{if} = 1 - w_{ii} - \sum_{n=1}^{\mathbb{N}-1} w_{in}.$$

We then have to define the investor's budget constraint, also maximize the utility subject to the budget constraint and impose regularity conditions to obtain the solutions. We choose a special case of iso-elastic utility function, $\lim_{\eta \to 1} \frac{Q^{1-\eta}-1}{1-\eta} = \ln(Q)$.[41] With the logarithmic utility, we obtain that consumption is

$$c = \rho I$$

and optimal shares are

$$w_{ii} = \sigma_P^{-2} \left[ (R_P - r_f) - w_{if}' \omega \right]$$

and

$$w_{in} = \tilde{\Sigma}_F^{-1} \left[ (R_F - r_f \iota) - \omega w_{ii} - \boldsymbol{S}^{-1}(w_{in}) \Pr(q_{in}) \right]$$
$$= \tilde{\Sigma}_F^{-1} \left[ (R_F - r_f \iota) - \omega \sigma_P^{-2} \left[ (R_P - r_f) - w_{if}' \omega \right] - \boldsymbol{S}^{-1}(w_{in}) \Pr(q_{in}) \right],$$

where $\Pr(q_{in})$ is the probability of contract violation in country $n$ for investment from $i$, $\text{diag}(w_{in})$ is a diagonal matrix of weights and

$$\boldsymbol{S}(w_{in}) = I_{\mathbb{N}-1} - \text{diag}(w_{in}) = \begin{bmatrix} 1 - w_{i1} & 0 & \cdots & 0 \\ 0 & 1 - w_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 - w_{i\mathbb{N}} \end{bmatrix}.$$

These equations show that the international investor behaves as a mean-variance investor. For example, the share of capital invested in each foreign economy equals its expected excess return relative to its variance and covariance and corrected for the risk of expropriation.

---

[40] See Tille and van Wincoop (2010) which focus on the new approximation and solution method to compute optimal shares in the DSGE framework.

[41] We have not used the function form of utility function in the main text, as FDI has been driven by supply side arguments. To introduce portfolio flows with partial equilibrium flavor, we analyze demand side and its implications for investment.

With this in mind, one can derive the optimal investment shares,

$$w_{ni} = \gamma_{ni} \left( \tau_{ni} \right) \left[ \mathbb{E} \left( R_i - R_f \right) - \frac{\theta_n}{\theta_g} \mathbb{E} \left( R_g - R_f \right) \right],$$

where $\gamma_{ni} \left( \tau_{ni} \right)$ is a constant that includes $\tau_{ni}$, $\theta_n / \theta_g$ determines the strength of a transmitter of global payoff innovation with $\theta_g = \sum_{n=1}^{\mathbb{N}} \left( I_n / I \right) \theta_n$. Then, the optimal share can be re-written as

$$w_{ni} = \gamma_{ni} \left( \tau_{ni} \right) \widetilde{R}_i,$$

where $\widetilde{R}_i \equiv \mathbb{E} \left( R_i - R_f \right) - \left( \theta_i / \theta_g \right) \mathbb{E} \left( R_g - R_f \right)$. This yields the total equity holdings $\mathcal{E}_i$ by agents from country $i$ as

$$\mathcal{E}_i = \sum_{n=1}^{\mathbb{N}} w_{ni} W_i = \sum_{n=1}^{\mathbb{N}} \gamma_{ni} \left( \tau_{ni} \right) \widetilde{R}_i W_i,$$

which gives

$$W_i = \frac{\mathcal{E}_i}{\sum_{n=1}^{\mathbb{N}} \gamma_{ni} \left( \tau_{ni} \right) \widetilde{R}_i} = \mathcal{E}_i / \mathcal{P}_i,$$

where $\mathcal{P}_i \equiv \sum_{n=1}^{\mathbb{N}} w_{ni}$. This produces the total equity claim by country $i$ on country $n$ as

$$P_n^I I_{ni} = w_{ni} W_i = \gamma_{ni} \left( \tau_{ni} \right) \widetilde{R}_i \mathcal{E}_i / \mathcal{P}_i. \tag{A.2}$$

Bilateral asset demand depends on the price (risk-return ratio) of country $i$ equity relative to an overall price index. To close the model, we impose a set of market clearing equations,

$$\sum_{i=1}^{\mathbb{N}} P_n^I I_{ni} = P_n^I I_n = \mathcal{E}_n,$$

where $P_n^I I_n$ is the country $n$ equity supply which is equal to demand $\mathcal{E}_n$ in equilibrium. This produces

$$\sum_{i=1}^{\mathbb{N}} \frac{P_n^I I_n}{\mathcal{E}} = \sum_{i=1}^{\mathbb{N}} \frac{\gamma_{ni} \left( \tau_{ni} \right) \widetilde{R}_i \mathcal{E}_i}{\mathcal{E} \mathcal{P}_i} = \frac{\mathcal{E}_n}{\mathcal{E}},$$

therefore, the returns are described by

$$\widetilde{R}_i = \frac{\mathcal{E}_n}{\mathcal{E}} \frac{1}{\sum_{i=1}^{\mathbb{N}} \frac{\gamma_{ni}(\tau_{ni})\mathcal{E}_i}{\mathcal{E}\mathcal{P}_i}} = \frac{\mathcal{E}_j}{\mathcal{E}} \frac{1}{\Pi_n},$$

where $\Pi_n \equiv \sum_{i=1}^{\mathbb{N}} \frac{\gamma_{ni}(\tau_{ni})\mathcal{E}_i}{\mathcal{E}\mathcal{P}_i}$ is the so-called multilateral resistance variable that measures the average financial frictions for country $n$ as a destination country. Substituting this solution back into (A.2), we get the following gravity specification for bilateral asset holdings:

$$P_n^I I_{ni} = w_{ni} W_i = \frac{\mathcal{E}_n \mathcal{E}_i}{\mathcal{E}} \frac{\gamma_{ni} \left( \tau_{ni} \right)}{\Pi_n \mathcal{P}_i},$$

where optimal weights are $n$'s share in total equity weighted by the relative financial friction $\gamma_{ni} \left( \tau_{ni} \right) / \Pi_n$, i.e., $w_{ni} = \frac{\mathcal{E}_n}{\mathcal{E}} \frac{\gamma_{ni}(\tau_{ni})}{\Pi_n}$.

Introduction of the fixed costs of holding foreign assets partitions investors into two groups: investors with access to international financial markets (all equity) and investors with no such access (only domestic equity). Both groups can still invest in the risk-free and global assets. This changes the optimal shares for domestic investors, namely,

$$w_{nn}^D = \gamma_{nn} \left( \tau_{nn} \right) \widetilde{R}_n,$$

and investors with the access to international finance,

$$w_{ni}^A = \gamma_{ni}\left(\tau_{ni}\right)\widetilde{R}_i,$$

where $\widetilde{R}_i \equiv \mathbb{E}\left(R_i - R_f\right) - \left(\theta_i/\theta_g\right)\mathbb{E}\left(R_g - R_f\right)$. The market equilibrium yields

$$\mathcal{E}_i = \sum_{n=1}^{\mathbb{N}} w_{ni}^A W_n^A + w_{ii}^D W_i^D = \sum_{n=1}^{\mathbb{N}} \gamma_{ni}\left(\tau_{ni}\right)\widetilde{R}_i W_n^A + \gamma_{ii}\left(\tau_{ii}\right)\widetilde{R}_i W_i^D \tag{A.3}$$

and

$$\widetilde{R}_i = \frac{\mathcal{E}_i}{W\left[\sum_{n=1}^{\mathbb{N}} \gamma_{ni}(\tau_{ni})(W_n^A/W)+\gamma_{ii}(\tau_{ii})\left(W_i^D/W\right)\right]}$$
$$= \frac{\mathcal{E}_i}{W\left[\sum_{n=1}^{\mathbb{N}} \gamma_{ni}(\tau_{ni})(W_n^A/W)\delta_{in}\right]} = \frac{\mathcal{E}_i}{\Pi_i W},$$

where

$$\delta_{in} = \begin{cases} 1, & i \neq n \\ W_i/W_i^A, & i = n \end{cases}, \tag{A.4}$$

and $W_i = W_i^A + W_i^D$, $\Pi_i \equiv \sum_{n=1}^{\mathbb{N}} \gamma_{ni}\left(\tau_{ni}\right)\left(W_n^A/W\right)\delta_{in}$, $W = \sum_{n=1}^{\mathbb{N}} W_n^A$. Using the aggregate demand for equity in country $i$ in (A.3) yields

$$\mathcal{E}_i/\mathcal{P}_i = W_i^A,$$

hence,

$$\mathcal{P}_i = \frac{\mathcal{E}_i}{W_i^A} = \frac{\sum_{n=1}^{\mathbb{N}} \gamma_{ni}(\tau_{ni})\widetilde{R}_i W_n^A + \gamma_{ii}(\tau_{ii})\widetilde{R}_i W_i^D}{W_i^A} = \frac{\mathcal{E}_i}{\Pi_i W}\frac{\sum_{n=1}^{\mathbb{N}} \gamma_{ni}(\tau_{ni})W_n^A\delta_{in}}{W_i^A} = \frac{\mathcal{P}_i}{\Pi_i W}\sum_{n=1}^{\mathbb{N}} \gamma_{ni}\left(\tau_{ni}\right)W_n^A\delta_{in}$$
$$= \sum_{n=1}^{\mathbb{N}} \frac{\gamma_{ni}(\tau_{ni})\delta_{in}W}{\mathcal{E}_i\Pi_i} = \frac{1}{\mathcal{E}_i}\frac{\sum_{n=1}^{\mathbb{N}} \gamma_{ni}(\tau_{ni})\delta_{in}W}{\sum_{n=1}^{\mathbb{N}} \gamma_{ni}(\tau_{ni})(W_n^A/W)\delta_{in}} = \frac{1}{\mathcal{E}_i}\sum_{n=1}^{\mathbb{N}} \gamma_{ni}\left(\tau_{ni}\right)\delta_{in}W.$$

Plugging this result into flows equation, we obtain

$$M_{ni} = w_{ni}^A W_i^A = \gamma_{ni}\left(\tau_{ni}\right)\widetilde{R}_i\mathcal{E}_i/\mathcal{P}_i = \gamma_{ni}\left(\tau_{ni}\right)\frac{\mathcal{E}_n W_i^A}{W\Pi_n}$$
$$= \gamma_{ni}\left(\tau_{ni}\right)\frac{\mathcal{E}_n W_i^A\delta_{in}}{W\Pi_n} = \gamma_{ni}\left(\tau_{ni}\right)\frac{\mathcal{E}_n\mathcal{E}_i}{W\Pi_n\mathcal{P}_i}\delta_{in} \tag{A.5}$$

because $\delta_{in} = 1$ when $i \neq n$ and $W_i^A = \delta_{ii}W_i$. This is the equation reported in the main text (see (5.4)), and used for estimation.

## B  Econometric Methodology

### B.1  Estimating Relations

We first need to connect decisions made by firms and agents into one estimable setting. The main relations are given by

$$\begin{cases} M_{ni}^{FDI} = \left(1 - \omega_i - \omega_n + \omega_i\omega_n\right)v_n^{1-\sigma}V_{in}^{FDI}\frac{E_i E_n}{[P_n\Pi_i]^{1-\sigma}}, \\ M_{ni}^I = \omega_n\omega_i\gamma_{ni}\left(\tau_{ni}\right)\frac{E_i E_n}{W\Pi_n\mathcal{P}_i}\delta_{in}, \\ \frac{M_{ni}^I}{M_{ni}^{FDI}} = \frac{\omega_n\omega_i}{1-\omega_i-\omega_n+\omega_i\omega_n}\frac{\gamma_{ni}(\tau_{ni})}{v_n^{1-\sigma}}\frac{[P_n\Pi_i]^{1-\sigma}}{W\Pi_n\mathcal{P}_i}\frac{\delta_{in}}{V_{in}^{FDI}}. \end{cases}$$

These results follow from

$$M_{ni}^{FDI} = v_n^{1-\sigma}V_{in}^{FDI}\frac{(E_n - \mathcal{E}_n)(E_n - \mathcal{E}_n)}{[P_n\Pi_i]^{1-\sigma}}$$
$$= v_n^{1-\sigma}V_{in}^{FDI}\frac{E_n E_i}{[P_n\Pi_i]^{1-\sigma}}\left(1 - \left(\frac{\mathcal{E}_i}{E_i} + \frac{\mathcal{E}_n}{E_n} - \frac{\mathcal{E}_n\mathcal{E}_i}{E_n E_i}\right)\right),$$

when we introduce $\omega_i \equiv \mathcal{E}_i / E_i$ and $\sum_{n=1}^{N} M_{in}^{FDI} = \tilde{E}_i = E_n - \mathcal{E}_n$. When considered as a sum, $M_{ni}^{I} + M_{ni}^{FDI}$, total (bilateral) capital flow leads to the following cases:

$$
\begin{cases}
= \left[ \frac{v_n^{1-\sigma} V_{in}^{FDI}}{[P_n \Pi_i]^{1-\sigma}} \left(1 - (\omega_i + \omega_n - \omega_n \omega_i)\right) + \frac{\gamma_{ni}(\tau_{ni})\delta_{in}\omega_n\omega_i}{W \Pi_n \mathcal{P}_i} \right] E_n E_i, & \omega_i, \omega_n \neq 0, \\[2ex]
= \frac{v_n^{1-\sigma} V_{in}^{FDI} E_n E_i}{[P_n \Pi_i]^{1-\sigma}} (1 - \omega_i) = \frac{v_n^{1-\sigma} V_{in}^{FDI} E_n \tilde{E}_i}{[P_n \Pi_i]^{1-\sigma}} = \frac{v_n^{1-\sigma} V_{in}^{FDI} \tilde{E}_n \tilde{E}_i}{[P_n \Pi_i]^{1-\sigma}}, & \omega_i \neq 0, \omega_n = 0, \\[2ex]
= \frac{v_n^{1-\sigma} V_{in}^{FDI} E_n E_i}{[P_n \Pi_i]^{1-\sigma}} (1 - \omega_n) = \frac{v_n^{1-\sigma} V_{in}^{FDI} \tilde{E}_n E_i}{[P_n \Pi_i]^{1-\sigma}} = \frac{v_n^{1-\sigma} V_{in}^{FDI} \tilde{E}_n \tilde{E}_i}{[P_n \Pi_i]^{1-\sigma}}, & \omega_i = 0, \omega_n \neq 0, \\[2ex]
= \frac{v_n^{1-\sigma} V_{in}^{FDI} E_n E_i}{[P_n \Pi_i]^{1-\sigma}} = \frac{v_n^{1-\sigma} V_{in}^{FDI} \tilde{E}_n \tilde{E}_i}{[P_n \Pi_i]^{1-\sigma}}, & \omega_i, \omega_n = 0.
\end{cases}
$$

## B.2 Probit Specification

We start with the bilateral FDI flow. The log-linearized version of $Z_{in}$ (see footnote [20]) is

$$
\begin{aligned}
\mathsf{z}_{in} = (\sigma - 1)\ln(\sigma - 1) - \sigma \ln \sigma - \sigma \ln v_n - (\sigma - 1)\ln \varphi_L + (\sigma - 1)\mathsf{p}_n + \mathsf{e}_n \\
-\phi_i - \phi_n - \kappa \phi_{ni} + v_{ni} = \gamma_0 - \phi_i + \zeta_n - \sigma \ln v_n - \kappa \phi_{ni} + \eta_{ni},
\end{aligned} \tag{B.1}
$$

where $f_{in}^{FDI} = \exp(\phi_i + \phi_n + \kappa \phi_{ni} - v_{ni})$ is stochastic fixed investment cost, $\eta_{ni} \equiv u_{ni} + v_{ni} \sim N\left(0, \sigma_u^2 + \sigma_v^2\right)$ is i.i.d. (yet correlated with the error term $u_{in}$ in the gravity equation at intensive margin), $\phi_i$ is an investor's fixed effect, and $\zeta_n = (\sigma - 1)p_n + e_n - \phi_n$ is a destination country's fixed effect. These fixed effects are expected to control for the multilateral resistance terms. Note that $\mathsf{z}_{ni} > 0$ when $i$ invests to $n$, and $\mathsf{z}_{ni} = 0$ when it does not. Scaling $\mathsf{z}_{in}$ by the standard deviation $\sigma_{u+v}$, we obtain the probability $\rho_{ji}$ that $i$ decides invest to $j$, conditional on the observed variables,

$$
\begin{aligned}
\rho_{ni} &= \Pr\left(\mathsf{z}_{ni}^\star > 0 \mid \text{observed variables}\right) \\
&= \Phi\left(\gamma_0^\star + \phi_i^\star + \zeta_n^\star - \sigma^\star \ln v_n - \kappa^\star \phi_{ni}\right),
\end{aligned} \tag{B.2}
$$

where the starred variables refer to original ones divided by $\sigma_{u+v}$ and $\phi_i^\star$, $\xi_n^\star$ denote the fixed effects.[42] Normalization is justified since the coefficients of a probit can only be estimated up to a scale. Moreover, the Probit equation can be used to derive consistent estimates of $W_{ni}$ (which is crucial to avoid an omitted variable problem that biases the estimates). Further, note that inverse Normal maps estimated probability to the latent variable $\mathsf{z}_{ni}$, $\hat{\mathsf{z}}_{ni}^\star = \Phi^{-1}(\hat{\rho}_{ni})$. Note that $\eta_{ni}/\sigma_{u+v}$ has a unit normal distribution, therefore, in principle, a consistent estimate is obtainable from the inverse Mills ratio (selection hazard). We can generate it from the estimation of a probit model where the error term follows a standard normal distribution. Hence, the approximate estimate of $\mathsf{E}\left(\eta_{ni}^\star \mid \cdot, \mathsf{z}_{ni}^\star > 0\right)$ is simply $\phi\left(\hat{\mathsf{z}}_{ni}^\star\right)/\Phi\left(\hat{\mathsf{z}}_{ni}^\star\right)$. The estimate $\ln\left\{\exp\left(\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma - 1}\left(\hat{\mathsf{z}}_{ni}^\star + \mathsf{E}\left(\eta_{ni}^\star \widehat{\mid \cdot, \mathsf{z}_{ni}^\star > 0}\right)\right)\right) - 1\right\}$[43] helps controlling for the potential correlation between the investment frictions and the average productivity of firms in country $i$ which decide to invest to $n$ (note that investment barriers affect cutoff productivity level). However, this method works only under homoskedasticity, a requirement rarely met in practice with bilateral flows data. As a result, Silva and Tenreyro (2009) claim that conditional expectation of $\frac{\eta_{ni}}{\sigma_{u+v}} \equiv \frac{u_{ni} + v_{ni}}{\sigma_{u+v}}$, $\mathsf{E}\left(\eta_{ni}^\star \mid \cdot, \mathsf{z}_{ni}^\star > 0\right)$ is inappropriately corrected.[44] As is obvious from (B.3), $\eta_{ni}^\star$ is replaced by the conditional expectation

---

[42] In empirical trade gravity literature, fixed effects are interpreted as controlling for multilateral resistance terms (which embody trade costs across trading partners). Whereas in the regional science literature, the fixed effects are treated as the "repulsive" and "attractive" forces of the exporter and the importer, see, for example, Roy (2004). In this application, fixed effects control for the system-wide barriers of investment as embodied in (4.6) and (4.8).

[43] The selection equation has been derived from a firm-level decision, and it therefore does not contain the unobserved and endogenous variable $W_{ni} \equiv \max\left\{ Z_{ni}^{\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}} - 1, 0 \right\}$ that is related to the fraction of multinationals (direct investors). The estimate allows mapping $\mathsf{E}\left(w_{ni}^\star \mid \cdot, \mathsf{z}_{ni}^\star > 0\right)$ into the reported expression.

[44] Notably, $\eta_{ni}^\star$ enters the model non-linearly that further complicates analysis - an estimator that is robust to incidental distributional assumptions is non-existent. More importantly, since $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ depends on $\sigma_u$ and $\sigma_v$, if either of the random components of the model is heteroskedastic, this makes the whole term a function of the regressors, which will then enter the model in a much more complex form.

and parameter $\beta_{u\eta} = \rho_{u,\eta}\frac{\sigma_u}{\sigma_v}$. The reason lies in Jensen's inequality, i.e., for any non-linear function $f(\cdot)$, $f\left(\mathsf{E}\left(\eta_{ni}^{\star} \mid \cdot, z_{ni}^{\star} > 0\right)\right) = f\left(\phi\left(z_{ni}^{\star}\right)/\Phi\left(z_{ni}^{\star}\right)\right) \neq \mathsf{E}\left(f\left(\eta_{ni}^{\star}\right) \mid \cdot, z_{ni}^{\star} > 0\right)$. Therefore, $f\left(\phi\left(\hat{z}_{ni}^{\star}\right)/\Phi\left(\hat{z}_{ni}^{\star}\right)\right)$ is not a consistent estimator. However, the approximation is justified if deviation form the conditional mean, $\eta_{ni}^{\star} - \phi\left(z_{ni}^{\star}\right)/\Phi\left(z_{ni}^{\star}\right)$, is linear in the random term because it would be absorbed by the error of the equation. Notably, this approximation is likely to be reasonably accurate in many empirical studies.

Fortunately, it is possible to test for heteroskedasticity in the probit specification. Note that $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ is proportional to the standard deviation of the error in (B.2). Hence, adding additional regressors $z_{ni}^{\star 2}$ and $z_{ni}^{\star 3}$ to the probit specification and checking for the joint significance would resemble a special case of White's test for heteroskedasticity. The test is analogous to a two-degrees-of-freedom RESET test (Ramsey, 1969), is also particularly interesting in that it can be interpreted as a normality test (see Newey, 1985). Therefore, this simple test provides a direct check for the validity of the main distributional assumptions required for consistent estimation of the model of interest. RESET test (based on a linear index) for the probit model should first be applied for the the validity of the results obtained with the two-stage estimators.[45]

Substituting estimates into the FDI gravity equation leads to

$$m_{ni} = \boldsymbol{\beta}'\boldsymbol{X} + \ln\left\{\exp\left(\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}\left(\hat{z}_{ni}^{\star} + \mathsf{E}\left(\widehat{\eta_{ni}^{\star} \mid \cdot, z_{ni}^{\star}} > 0\right)\right)\right) - 1\right\} + \beta_{u\eta}\mathsf{E}\left(\widehat{\eta_{ni}^{\star} \mid \cdot, z_{ni}^{\star}} > 0\right) + e_{ni},$$
(B.3)

where $\mathsf{E}\left(u_{ni} \mid \cdot, z_{ni}^{\star} > 0\right) = \beta_{u\eta}\mathsf{E}\left(\eta_{ni}^{\star} \mid \cdot, z_{ni}^{\star} > 0\right)$, $\beta_{u\eta} = \rho_{u,\eta}\frac{\sigma_u}{\sigma_v}$ and $e_{in}$ is an *iid* error term $\mathsf{E}\left(e_{ni} \mid \cdot, z_{ni}^{\star} > 0\right) = 0$. Note that the term $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ contains a relationship between Pareto shape parameter $k$ which is directly related to firm heterogeneity and as such affects the variance of the error term for different values of $k$.[46] It is also related to the measure of the extensive margin, which is associated with the elasticity of substitution $\sigma$.[47] Further, when $\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}$ increases, the following approximation improves:

$$\ln\left\{\exp\left(\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}\left(\hat{z}_{ni}^{\star} + \mathsf{E}\left(\widehat{\eta_{ni}^{\star} \mid \cdot, z_{ni}^{\star}} > 0\right)\right)\right) - 1\right\} \approx \frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}\left(\hat{z}_{ni}^{\star} + \mathsf{E}\left(\widehat{\eta_{ni}^{\star} \mid \cdot, z_{ni}^{\star}} > 0\right)\right).$$

Hence, in this case the terms enter (B.3) approximately linearly. Helpman, Melitz, and Rubinstein (2008) estimate (B.3) using nonlinear least squares as parameters enter nonlinearly in the expression for $w_{ni}^{\star}$. The use of $\mathsf{E}\left(\eta_{ni}^{\star} \mid \cdot, z_{ni}^{\star} > 0\right)$ to control for $\mathsf{E}\left(u_{ni} \mid \cdot, z_{ni}^{\star} > 0\right)$ is the Heckman (1979) correction for sample selection. Notice, however, that the unobserved firm-level heterogeneity is corrected by the additional control $\hat{z}_{ni}^{\star}$. At first, (B.1) is estimated, then proceeded to (B.3). To avoid reliance on the normality assumption for the unobserved capital frictions, valid excluded variables for the second stage are needed. Observe that $\hat{z}_{ni}^{\star}$ enters both (B.2) and (B.3). The exogenous source of variation that enters only (B.2) and helps breaking collinearity problem is $\phi_{ni}$. Economically, this exclusion restriction concerns fixed investment costs which have bearance on variable transaction costs. In other words, we need variables which correlate with the $z_{ni}$ but not with the residual of the second stage equation. Further, observe that $z_{ni}^{\star}$ includes only destination fixed costs which can be fruitfully exploited to estimate FDI gravity.

Note that the statistical implications of the log transformation of multiplicative models was well understood at least from Teeken and Koerts (1972). Recently, Jia and Rathi (2008) proposed an anti-log

---

[45] Presence of heteroskedasticity highly complicates the identification of the effects of the covariates in the intensive and extensive margins as pointed out by Silva and Tenreyro (2009).

[46] In particular, a greater dispersion of productivity (a lower $k$) decreases the concentration of firms around the cutoff. This implies a smaller decrease in the mass of operating firms for a given increase in the productivity cutoff.

[47] The change in $W_{ni}$, as defined in Footnote 43, with respect to $k$ is always positive, though it is not straightforward to establish the speed of change, $\frac{\partial W_{ni}}{\partial k} = \frac{\sigma_{u+v}^2(k-\sigma+1)}{(\sigma-1)^2}Z_{ni}^{\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}}\ln Z_{ni} > 0$ since $Z_{ni} > 1$ and $\frac{\partial^2 W_{ni}}{\partial k^2} = \frac{\sigma_{u+v}^2}{(\sigma-1)^2}Z_{ni}^{\frac{\sigma_{u+v}(k-\sigma+1)}{\sigma-1}}\ln Z_{ni}\left(1 + \frac{\sigma_{u+v}(k-\sigma+1)}{(\sigma-1)}\ln Z_{ni}\right)$ where cutoff productivities were assumed independent of $k$. The second derivative is negative whenever $\sigma_{u+v} < -\frac{\sigma-1}{k-\sigma+1}(\ln Z_{ni})^{-1}$. Only in this case larger $k$ (or lower firm heterogeneity, see Chaney, 2008) induces an increase (at a diminishing rate) of extensive margin.

transformation which reduces the biases of exponentiation in the models with heteroskedastic errors.[48]

### B.2.1  Pseudo Poisson Maximum Likelihood estimator

The Pseudo Poisson Maximum Likelihood estimator (PPML), put forward by Santos-Silva and Tenreyro (2007), is consistent even if the data does not follow a Poisson distribution. Heteroskedasticity undermines the use of log-linearization by making conditional expectation dependent on the regressors, and Jensen's inequality points to inconsistency recovering estimated flows to levels. Given there is no information on the pattern of heteroskedasticity, PPML seems natural in that it weighs all the observations equally, since the expected value is assumed to be proportional to the variance, $\mathsf{E}\left(M_{ni}\right) \propto Var\left(M_{ni}\right)$. Since this assumption is unlikely to hold in practice, a failure to fully account for the heteroskedasticity should be corrected for by an Eicker-White (Eicker, 1963; White, 1980) robust covariance matrix estimator.

There are also other reasons to apply the Poisson quasi-maximum likelihood (QML) estimator. As noted by Santos-Silva and Tenreyro (2007), least squares potentially introduce biases as they drop observations for which the reported flow value is zero. Moreover, inconsistency results in cases of heteroskedastic disturbances with the variance dependence on one or more of the regressors. Both concerns can be addressed using Poisson QML with a robust error covariance matrix. Recently, Arvis and Shepherd (2011) demonstrated that Poisson QML estimator is the only in the class of (quasi-) ML estimators to preserve total flows between the actual and estimated bilateral trade matrices. Coined as the "adding up" problem, it is an important issue as in our application FDI flow equation is linear in a constant term with an elasticity of 1, $\varepsilon_{\hat{M}_{int},\,\mathrm{const}} \equiv \partial \ln \hat{M}_{int}/\partial \ln \mathrm{const} = 1$ or $\partial \hat{M}_{int}/\partial \mathrm{const} = (1/\mathrm{const})\,\hat{M}_{int}$. Maximizing the Poisson QML on the constant scale equalizes the total values of foreign investment. However, its unintended consequence is a perfect fit between the fixed effects and the multilateral resistance terms (see Head and Mayer (2015) and references therein).

Table B.1 reports the estimates of Poisson regression whose specification is $\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in},\, M_{in} > 0\right) = \exp\left(\boldsymbol{X}_{in}\beta\right)$ with conditional expectation of error being zero. This approach should control for heteroskedasticity and zero values of capital flows, however, lacks mechanisms that give rise to zeroes, also does not model cross-sectional dependencies or unobserved factors. Regarding results, they are similar compared to the baseline ones once directional fixed effects are included: destination income enters negatively (though insignificantly), source income exerts positive sign as does labor productivity, and lending rate is of negative sign (though marginally insignificant). Portfolio regressions mainly differ in the sign of contract risk - it is always negative as opposed to the baseline results. However, once we allow for flexible dependence between the two decisions - in the robustness checks that follow - we again arrive at the similar conclusion with higher risks being associated with the decision not to invest.

### B.3  Tobit Specification

As an extension, we also consider a tobit specification. It is especially suited for portfolio investment when we allow for investor's decisions at both margins to be correlated. In other words, we introduce a possibility for the intensive and extensive margins, after controlling for covariates, to be driven by common factors.

---

[48] With the error term normality, $\mathsf{m}_{ni} = \mu_n + \boldsymbol{\beta}'\boldsymbol{x}_{ni} + \varepsilon_{ni}$ where $\mathsf{m}_{ni} \equiv \ln M_{ni}$ and $\varepsilon_{ni} \sim N\left(0, \sigma_n^2\right)$, we have $\mathsf{E}\left(\mathsf{m}_{ni}\right) = \exp\left(\tilde{\mu}_n + \tilde{\boldsymbol{\beta}}'\boldsymbol{x}_{ni} + \frac{\tilde{\sigma}_n^2}{2}\right)$. Note that the variables with tilde must be consistent for $\mathsf{E}\left(\mathsf{m}_{ni}\right)$ to be also consistent. However, this expression over-estimates the expected value because of convexity of exponential function. However, ignoring correlations between $\tilde{\mu}_n$, $\tilde{\beta}_n$ and $\tilde{\sigma}_n^2$, one can write: $\mathsf{E}\left(\mathsf{m}_{ni}\right) = \exp\left(\tilde{\mu}_n\right)\prod_{n=0}^{I}\exp\tilde{\beta}_n\boldsymbol{x}_{ni}\exp\left(\frac{\tilde{\sigma}_n^2}{2}\right)$ where $\mathsf{E}\exp\left(\tilde{\mu}_n\right) = \exp\left(\mu_n + \frac{\sigma_{\tilde{\mu}_n}^2}{2}\right)$ and $\mathsf{E}\exp\left(\tilde{\beta}_n\boldsymbol{x}_{ni}\right) = \exp\left(\beta_n\boldsymbol{x}_{ni} + \frac{\left(x_{ni}\tilde{\sigma}\beta_n\right)^2}{2}\right)$, therefore, $\widetilde{\mathsf{E}\left(\mathsf{m}_{ni}\right)} = \exp\left(\tilde{\mu}_n - \frac{\tilde{\sigma}_{\mu_n}^2}{2}\right)\prod_{n=0}^{I}\exp\left(\tilde{\beta}_n\boldsymbol{x}_{ni} - \frac{\left(x_{ni}\tilde{\sigma}\beta_n\right)^2}{2}\right)\exp\left(\frac{\tilde{\sigma}_n^2}{2}\right)$, where the adjustment for $\exp\left(\tilde{\sigma}_n^2/2\right)$ is omitted for simplicity.

Tab. B.1: Poisson Regressions for Capital Flows

| FDI | LIC | | MIC | | Overall Sample | |
|---|---|---|---|---|---|---|
| Variables | Without DFE | With DFE | Without DFE | With DFE | Without DFE | With DFE |
| Log Destination | 0.5767*** | -0.5724 | 0.7733*** | -0.0550 | 0.7564*** | -0.0989 |
| Income | (0.0756) | (2.0248) | (0.0315) | (0.9153) | (0.0244) | (0.7901) |
| Log Source | 0.7112*** | 1.8226 | 0.5853*** | 2.1060* | 0.5909*** | 2.1778** |
| Income | (0.1092) | (2.4859) | (0.0285 ) | (1.1582) | (0.0290) | (1.0234) |
| Log Lending | -1.6749*** | -0.8244 | 0.2198*** | -0.3689 | 0.1259* | -0.3819* |
| Rate | (0.3435) | (0.5037) | (0.0647) | (0.2329) | (0.0667) | (0.2219) |
| Log Labor | 0.1289* | 4.4833** | 0.6380*** | 1.6811 | 0.6341*** | 1.7355** |
| Productivity | (0.0702) | (2.0868) | (0.1024) | (1.0239) | (0.0702) | (0.8866) |
| Observations | 2476 | 2476 | 4616 | 4616 | 7092 | 7092 |
| *BIC* | 192333.6 | 66372.29 | 2181070 | 1531468 | 2457263 | 1621971 |

Note: time coverage is 2002-2010. Solved using Quasi ML Fisher scoring (iterated least-squares and expected information matrix). DFE stands for directional (source and destination countries) fixed effects. Standard errors are bootstrapped (400 replications). *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

| Portfolio | LIC | | MIC | | Overall Sample | |
|---|---|---|---|---|---|---|
| Variables | Without DFE | With DFE | Without DFE | With DFE | Without DFE | With DFE |
| Log Destination | 0.3444* | 1.0447* | 0.2003*** | 0.0844 | 0.2208*** | 0.0863 |
| Income | (0.1790) | (0.5564) | (0.0161) | (0.0636) | (0.0143) | (0.0608) |
| Log Source | -0.2111 | -0.9182 | -0.1146*** | 1.0059*** | -0.1172*** | 0.9993*** |
| Income | (0.2032) | (0.9054) | (0.0292) | (0.1399) | (0.0281) | (0.1398) |
| Contract | -1.3748** | -0.0769 | -0.4508*** | -0.4886** | -0.3601*** | -0.4886** |
| Risk | (0.5512) | (0.7988) | (0.0963) | (0.2133) | (0.0944) | (0.1962) |
| Observations | 212 | 212 | 2793 | 2793 | 3005 | 3005 |
| *BIC* | 21379.51 | 9762.037 | 7696402 | 1233491 | 7876161 | 1250267 |

Note: time coverage is 2002-2010. Solved using Quasi ML Fisher scoring (iterated least-squares and expected information matrix). DFE stands for directional (source and destination countries) fixed effects. Standard errors are bootstrapped (400 replications). *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

However, the standard tobit does not suit as it restricts the selection mechanism to be from the same model as that generating the outcome variable. Moreover, failure of homoskedasticity and Normality in the tobit model has serious consequences. The part models are more flexible and less prone to deviations from these assumptions.[49] The statistical problem, acknowledged by Felbermayr and Kohler (2006) in their corner-solutions version of the gravity equation, is that the conditional mean of actual flow $\mathsf{m}_{in}$ cannot be linear in explanatory variables, because there is a positive probability mass at $\mathsf{m}_{in} = 0$. One of the possible solutions - nonlinear least squares (NLS) - are imperfect because of potential heteroskedasticity (corner outcomes) in which case NLS are inefficient. More importantly, the coefficients obtained by NLS estimation of a model for $\mathsf{E}\left(\ln M_{in} \mid \cdot\right)$ are hard to give an adequate interpretation. NLS do not allow us to empirically identify the intensive and extensive margins of world trade.

Let's first start with the observation that the observed capital flows can be treated as continuously distributed, non-negative latent variable. Given the fixed costs of entering foreign markets, the decision of investing will be determined by the latent variables $Z_{in}$ or $\delta_{in}$. Let's first start with the assumption employed by Felbermayr and Kohler (2006) that intensive and extensive margins are independent conditional on explanatory variables, denoted by $D\left(M_{in}^{\star} \mid Z_{in}, \boldsymbol{X}_{in}\right) = D\left(M_{in}^{\star} \mid \boldsymbol{X}_{in}\right)$ where $D\left(\cdot \mid \cdot\right)$ is the conditional distribution, we end up with the two-part or hurdle model, see Wooldridge (2010).[50] A truncated normal hurdle model by Cragg (1971) with the latent variable (positive capital flows in our analysis) is assumed to have a truncated normal distribution. Since we are mainly interested in recovering $\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right)$, we may not make distributional assumptions. One method is the smearing estimate by Duan (1983). However, a direct method would be to estimate $\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right)$ as

$$\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right) = \exp\left(\boldsymbol{X}_{in}\beta\right),$$

and this contains $M_{in}^{\star} = \exp\left(\boldsymbol{X}_{in}\beta + u\right)$. A quasi-MLE is (Fisher) consistent given the conditional mean is correctly specified. Note that capital flows need not correspond to the chosen density, what only needed is the same range in the conditional mean as allowed in the chosen density of linear exponential family (LEF). For every QMLE in LEF there is an asymptotically (or $\sqrt{N}$) equivalent weighted non-linear least squares (WNLS) estimator if the conditional mean is well specified despite the conditional variance. Hence, we first estimate a probit $\Pr\left(M_{in} > 0 \mid \boldsymbol{X}_{in}\right) = \Phi\left(\boldsymbol{X}_{in}\gamma\right)$, then obtain QMLE estimates of $\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right) = \exp\left(\boldsymbol{X}_{in}\beta\right)$. Finally, we can estimate $\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in}\right) = \Phi\left(\boldsymbol{X}_{in}\gamma\right)\exp\left(\boldsymbol{X}_{in}\beta\right)$.

### B.3.1    Correlation Between The Margins

As mentioned, the assumption of independent mechanisms generating intensive and extensive margins seems to be theoretically unjustified. We employ the type II Tobit model which allows for common unobserved factors thereby extending type I Tobit model. Let $M_{in}^{\star} = \exp\left(\boldsymbol{X}_{in}\beta + u\right)$ and $\{\Pr\left(M_{in} > 0 \mid \boldsymbol{X}_{in}\right)\} = 1\left(\boldsymbol{X}_{in}\gamma + v > 0\right)$, and allow $u$ and $v$ be correlated. That is, let correlation be $\rho = Cov\left(u, v\right)/\sigma$ where $\sigma = \sqrt{Var\left(u\right)}$. Moreover, let $D\left(\ln\left(M_{in}^{\star}\right)\right) = N\left(\boldsymbol{X}_{in}\beta, \sigma^2\right)$. To find the density $f\left(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right)$, we, as in Wooldridge (2010), use the change-of-variables formula, leading to

$$f\left(M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right) = g\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right)/M_{in},$$

---

[49] Flexibility comes from the fact that tobit is just a probit model combined with a truncated regression model, with the coefficient vectors in the two models restricted to be proportional to each other. The second stage can be modeled using $\mathsf{E}\left(\ln M_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right)$ where the crucial assumption being linearity in $\boldsymbol{X}_{in}$ (alternatively, full maximum likelihood can be used for efficiency reasons but this requires that errors are jointly normally distributed). Given the interest concentrates on the conditional expectation only, the log of capital can be estimated as exponential using quasi-MLE without further distributional assumptions. Moreover, as the theoretical motivation purports, the set of variables affecting decision to invest and the volume are not identical. This makes results more robust as identification strategy does not rest entirely on distributional assumptions and the nonlinearity of the inverse of Mills ratio.

[50] A standard tobit model does not fit because it assumes that the two decisions are made jointly and affected by the same explanatory variables.

where $g\left(\cdot \mid \boldsymbol{X}_{in}, M_{in} > 0\right)$ is the conditional density of $M_{in}^{\star}$. By Bayesian rule,

$$\Pr\left(M_{in} > 0 \mid \boldsymbol{X}_{in}\right) \ g\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right) = \Pr\left(M_{in} > 0 \mid \mathsf{m}_{in}, \boldsymbol{X}_{in}\right) h\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}\right),$$

where $h\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}\right)$ is the conditional density of $\ln\left(M_{in}\right)$. Let $\mathbf{1}\left(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0\right) = \mathbf{1}\left(\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)\,u + e > 0\right)$ where $v = (\rho/\sigma)\,u + e$, $u = \mathsf{m}_{in} - \boldsymbol{X}_{in}\beta$ and $e \mid \boldsymbol{X}_{in}, u \sim N\left(0, 1 - \rho^2\right)$. This leads to

$$\Pr\left(M_{in} > 0 \mid \mathsf{m}_{in}, \boldsymbol{X}_{in}\right) = \frac{\Pr(M_{in} > 0 \mid \boldsymbol{X}_{in})g(\mathsf{m}_{in}\mid\boldsymbol{X}_{in},\,M_{in} > 0)}{h(\mathsf{m}_{in}\mid\boldsymbol{X}_{in})}$$
$$= \Phi\left(\left[\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)\left(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta\right)\right]\left(1 - \rho^2\right)^{-\frac{1}{2}}\right).$$

Recalling that $h\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}\right) \sim N\left(\boldsymbol{X}_{in}\beta, \sigma^2\right)$ leads to

$$f\left(M_{in} \mid \boldsymbol{X}_{in}\right)$$
$$= \Phi\left(\left[\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)\left(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta\right)\right]\left(1 - \rho^2\right)^{-\frac{1}{2}}\right)\phi\left(\left(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta\right)/\sigma M_{in}\right).$$

Combining with the density for $M_{in} = 0$, we obtain the log-likelihood

$$l_i\left(\theta\right) = 1\left[M_{in} = 0\right]\ln\left[1 - \Phi\left(\boldsymbol{X}_{in}\boldsymbol{\gamma}\right)\right] + 1\left[M_{in} > 0\right]$$
$$\times\left\{\ln\left[\Phi\left(\left[\boldsymbol{X}_{in}\boldsymbol{\gamma} + (\rho/\sigma)\left(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta\right)\right]\left(1 - \rho^2\right)^{-\frac{1}{2}}\right)\right] + \ln\left[\phi\left(\left(\mathsf{m}_{in} - \boldsymbol{X}_{in}\beta\right)/\sigma\right)\right] - \ln\sigma - \mathsf{m}_{in}\right\}.$$

Note that $\rho = 0$ yields a standard lognormal hurdle model. The challenge with this model, in which $(u, v)$ is independent of $\boldsymbol{X}_{in}$ and jointly normal, is the identification. Given the set of explanatory variables in $M_{in}^{\star} = \exp\left(\boldsymbol{X}_{in}\beta + u\right)$ are the same as in $\left\{\Pr\left(M_{in} > 0 \mid \boldsymbol{X}_{in}\right)\right\} = 1\left(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0\right)$. Note that the latter term is a function of $(\boldsymbol{X}_{in}, v)$, the law of iterated expectations yields

$$\mathsf{E}\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}, v\right) = \boldsymbol{X}_{in}\beta + \mathsf{E}\left(u \mid \boldsymbol{X}_{in}, v\right) = \boldsymbol{X}_{in}\beta + \mathsf{E}\left(u \mid v\right) = \boldsymbol{X}_{in}\beta + \rho\sigma v,$$
$$\mathsf{E}\left(v \mid \boldsymbol{X}_{in}, 1\left(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0\right) = 1\right) = \lambda\left(\boldsymbol{X}_{in}\boldsymbol{\gamma}\right) \equiv \frac{\phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})}{\Phi(\boldsymbol{X}_{in}\boldsymbol{\gamma})},$$
$$\mathsf{E}\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}, 1\left(\boldsymbol{X}_{in}\boldsymbol{\gamma} + v > 0\right) = 1\right) = \boldsymbol{X}_{in}\beta + \rho\sigma\lambda\left(\boldsymbol{X}_{in}\boldsymbol{\gamma}\right),$$
$$\mathsf{E}\left(\mathsf{m}_{in} \mid \boldsymbol{X}_{in}, M_{in} > 0\right) = \boldsymbol{X}_{in}\beta + \rho\sigma\lambda\left(\boldsymbol{X}_{in}\boldsymbol{\gamma}\right).$$

If $\boldsymbol{\gamma}$ can be consistently estimated, then the last equation can be nominally identified. However, the nonlinearity of $\lambda\left(\boldsymbol{X}_{in}\boldsymbol{\gamma}\right)$ acts as a poor identification strategy, especially over linear ranges. With unrestricted $\beta$, $\lambda\left(\boldsymbol{X}_{in}\boldsymbol{\gamma}\right)$ is a function of $\boldsymbol{X}_{in}$, arbitrarily close to linear if we relax a probit model on $\Pr\left(M_{in} > 0 \mid \boldsymbol{X}_{in}\right)$. If the identification is entirely based on $\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in}\right)$, the lognormal hurdle (note that $\rho\sigma$ is not separately identified by $\mathsf{E}\left(M_{in} \mid \boldsymbol{X}_{in}\right)$ since $\boldsymbol{X}_{in}$ includes a constant) and the ET2T models with the same set of regressors yield no differences. Hence, ET2T is more convincing when the covariates determining the participation decision strictly contain those affecting the amount decision. Then the model is

$$M_{in} = 1\left[\boldsymbol{X}_{in}\boldsymbol{\gamma} + v \geq 0\right] \times \exp\left(\tilde{\boldsymbol{X}}_{in}\tilde{\beta} + u\right),$$

where $\tilde{\boldsymbol{X}}_{in} \subset \boldsymbol{X}_{in}$ and both include unity as the first element. Given at least one exclusion restriction, $\tilde{\beta}$ and $\rho\sigma$ are better identified because $\lambda\left(\boldsymbol{X}_{in}\boldsymbol{\gamma}\right)$ is not an exact function of $\tilde{\boldsymbol{X}}_{in}$. Note that a linear rather than exponential model cannot be applied: linear type II Tobit model allows for negative outcomes on $M_{in}$.

## C Data Issues

### C.1 Country Coverage

Developing Countries are selected by PRGT-Eligibility. The list of countries included in the analysis is produced in the table below (71 countries in total):

| Afghanistan | Armenia | Bangladesh | Benin |
|---|---|---|---|
| Bhutan | Bolivia | Burkina Faso | Burundi |
| Cambodia | Cameroon | Cape Verde | Central African Rep. |
| Chad | Comoros | Congo, Dem. Rep. of | Congo, Republic of |
| Cï¿œte d'Ivoire | Djibouti | Dominica | Eritrea |
| Ethiopia | Gambia, The | Georgia | Ghana |
| Grenada | Guinea | Guinea-Bissau | Guyana |
| Haiti | Honduras | Kenya | Kiribati |
| Kyrgyz Republic | Lao People's Dem.Rep | Lesotho | Liberia |
| Madagascar | Malawi | Maldives | Mali |
| Mauritania | Moldova | Mongolia | Mozambique |
| Myanmar | Nepal | Nicaragua | Niger |
| Nigeria | Papua New Guinea | Rwanda | Samoa |
| Sï¿œo Tomï¿œ & Prï¿œencipe | Senegal | Sierra Leone | Solomon Islands |
| St. Lucia | St. Vincent & Grens. | Sudan | Tajikistan |
| Tanzania | Timor-Leste | Togo | Tonga |
| Uganda | Uzbekistan | Vanuatu | Vietnam |
| Yemen, Rep. | Zambia | Zimbabwe | |

## C.2   OECD Data

FDI Outflows are expressed in US dollars. The country coverage of reporting economies are (25 countries in total):

| Australia | Austria | Belgium | Canada |
|---|---|---|---|
| Denmark | Finland | France | Germany |
| Greece | Iceland | Ireland | Israel |
| Italy | Japan | Korea | Luxembourg |
| Netherlands | New Zealand | Norway | Portugal |
| Spain | Sweden | Switzerland | United Kingdom |
| United States of America | | | |

## C.3   CPIS Data

The CPIS measures the global stock and geographical distribution of portfolio investment holdings, as reported by creditor countries. The survey may have gaps in coverage owing to nonparticipation of some important investing countries and international financial centers, as well as difficulties faced by many participating countries in capturing cross-border portfolio investment by households (and in some cases, enterprises) that do not use the services of resident custodians. The stocks are measured at market value; thus, annual changes reflect valuation effects and flows.

## C.4  Other Data

Other variable used in the empirical exercises are collected in the table C.3, together with the sources of data.

Tab. C.3: Sources of remaining variables

| Variable | Data Source |
|---|---|
| Contract risk* | ICRG, The PRS Group, Inc. |
| Startup costs | World Bank's Starting Business Database |
| The marginal productivity of capital (MPK) | Lowe, Papageorgiou, and Perez-Sebastian (2012) |
| Aggregate income | IMF (WEO) |
| Wages | ILO (KLIM) |
| Lending rate | World Bank |
| Proxy for MPK (robustness check) | IMF (WEO) |
| Proxy for labour productivity (cost/GDP) | World Bank |
| * Contract Viability is the variable used, it measures the risk of unilateral contract modification, cancellation or outright expropriation. A higher value reflects lower risk. | |

# D  Robustness Checks

## D.1  Panels with MPK

Tab. D.1: Panel Results of Bilateral FDI Ignoring the Extensive Margin

|  | LIC | | | MIC | | | Overall Sample | | |
|---|---|---|---|---|---|---|---|---|---|
| Variables | Pooled OLS | FE | RE (GLS) | Pooled OLS | FE | RE (GLS) | Pooled OLS | FE | RE (GLS) |
| Log Destination | 0.84 | 1.25 | 0.86 | -1.18* | -1.08*** | -1.18** | -0.950* | -0.82*** | -0.96** |
| Income | (1.516) | (1.272) | (1.506) | (0.609) | (0.322) | (0.520) | (0.555) | (0.279) | (0.483) |
| Log Source | 2.49 | 2.61 | 2.53 | 5.25*** | 5.33*** | 5.25*** | 4.87*** | 4.92*** | 4.87*** |
| Income | (1.636) | (2.036) | (1.617) | (0.836) | (0.432) | (0.654) | (0.748) | (0.395) | (0.603) |
| Log MPK | 0.44 | 0.57 | 0.44 | -0.41 | -0.48 | -0.41 | -0.27 | -0.33 | -0.27 |
|  | (0.841) | (0.326) | (0.696) | (0.323) | (0.328) | (0.280) | (0.302) | (0.278) | (0.259) |
| Log Labor | 0.63 | -0.92 | 0.49 | 2.55*** | 2.37*** | 2.55*** | 2.46*** | 2.23*** | 2.47*** |
| Productivity | (2.545) | (1.538) | (2.710) | (0.690) | (0.385) | (0.615) | (0.657) | (0.302) | (0.590) |
| Observations | 499 | 499 | 499 | 2657 | 2657 | 2657 | 3156 | 3156 | 3156 |
| Within $R^2$ |  | 0.11 | 0.11 |  | 0.20 | 0.20 |  | 0.19 | 0.19 |
| Between $R^2$ |  |  | 0.61 |  |  | 0.69 |  |  | 0.70 |
| Overall $R^2$ |  |  | 0.46 |  |  | 0.59 |  |  | 0.60 |

Note: time coverage is 2002-2010. Pooled OLS refer to pooled least squares where any time-specific effects are assumed to be fixed and the individual affects are centered around a common intercept; cluster robust standard errors are reported, also directional fixed effects (source and destination countries) are included. FE stands for fixed effects with the reported Driscoll-Kraay standard errors which help controlling for spatial (cross-sectional) dependence. RE (GLS) for random effects generalized least squares estimator (effectively a weighted average of within and between estimators which are not reported separately), also directional fixed effects (source and destination countries) are included. Standard errors in parentheses; *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.

Tab. D.2: Bilateral FDI at Extensive and Intensive Margins with MPK

| Variables | LIC | | | MIC | | | Overall Sample | | |
|---|---|---|---|---|---|---|---|---|---|
| | Pooled Lognormal Hurdle | Pooled Exp. Type II Tobit | Mundlak Chamberlain Panel | Pooled Lognormal Hurdle | Pooled Exp. Type II Tobit | Mundlak Chamberlain Panel | Pooled Lognormal Hurdle | Pooled Exp. Type II Tobit | Mundlak Chamberlain Panel |
| **Intensive Margin of FDI** | | | | | | | | | |
| Log Destination Income | 1.20 | 0.654 | 1.04 | -1.90** | -1.67** | -1.19** | -1.59** | -1.32* | -0.795 |
| | (1.81) | (1.951) | (1.576) | (0.765) | (0.795) | (0.551) | (0.694) | (0.727) | (0.513) |
| Log Source Income | 1.23 | 1.045 | 1.46 | 5.25*** | 5.01*** | 5.34*** | 4.73*** | 4.38*** | 4.680*** |
| | (1.89) | (2.095) | (1.66) | (0.998) | (1.046) | (0.696) | (0.883) | (0.933) | (0.639) |
| Log MPK | 0.39 | -.684 | -.246 | -0.33 | -0.472 | 0.39 | -0.18 | -0.34 | 0.348 |
| | (0.91) | (1.029) | (0.802) | (0.37) | (0.403) | (0.441) | (0.347) | (0.373) | (0.376) |
| Log Labor productivity | 1.80 | 1.997 | 0.390 | 3.66*** | 3.15*** | 3.06*** | 3.57*** | 3.14*** | 2.83*** |
| | (3.31) | (3.850) | (3.344) | (0.89) | (0.971) | (0.717) | (0.849) | (0.940) | (0.683) |
| **Extensive Margin of FDI** | | | | | | | | | |
| Log Startup Costs | -0.0003 | -.0003 | -0.0003 | -0.0006 | -.0006 | -0.0006 | -0.0008 | -.0008 | -0.0008 |
| | (0.0009) | (0.0010) | (0.0009) | (0.0022) | (0.0022) | (0.0022) | (0.0008) | (0.0008) | (0.0008) |
| Log MPK | 7.090 | 6.966 | 7.090 | 10.36*** | 10.357*** | 10.36*** | 10.41*** | 10.36*** | 10.41*** |
| | (6.4386) | (6.342) | (6.4386) | (3.0960) | (3.091) | (3.0960) | (2.76) | (2.757) | (2.76) |
| Log Labor productivity | 0.0007*** | 0.0007*** | 0.0007*** | 0.00006*** | 0.00006*** | 0.00006*** | 0.00007*** | 0.00007*** | 0.00007*** |
| | (0.0002) | (0.0002) | (0.0002) | (0.00002) | (0.00002) | (0.00002) | (0.00002) | (.00002) | (0.00002) |
| $\hat{\sigma}$ ($\hat{\sigma_u}$ / $\hat{\sigma_\varepsilon}$) | 1.53 | 1.83 | .989/1.17 | 1.65 | 1.64 | 1.18/1.14 | 1.65 | 1.64 | 1.17/1.14 |
| | (0.052) | (0.177) | (0.091/0.046) | (0.027) | (0.028) | (0.047/0.018) | (0.025) | (.026) | (0.042/0.017) |
| $\hat{\rho}$ | - | -0.74*** | 0.418 | - | 0.075 | 0.52*** | - | 0.11* | 0.51*** |
| | | (0.131) | (0.052) | | (0.055) | (0.022) | | (0.06) | (0.02) |
| Observations (positive/all) | 499/1211 | 460/1483 | 458/1211 | 2657/4050 | 2447/4050 | 2447/4050 | 3156/5533 | 2907/5533 | 2907/5533 |
| Log likelihood | -3939.58 | -3860.76 | -3817.86 | -20994.15 | -20570.64 | -20126.67 | -25014.95 | -24514.76 | -24011.79 |

Note: time coverage is 2002-2010 The difference in observations in ET2T arises because all negative flows are excluded whereas in other estimations they are truncated as zeroes, potentially carrying information on the existence of a link but absence of a positive capital flow. Mundlak-Chamberlain method includes time averages of relevant explanatory variables and inverse Mills ratios from the first stage Probit regressions. Time averages are omitted in the first stage due to collinearity. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term and directional fixed effects (source and destination countries).

## D.2 Correlated Chamberlain-Mundlak Panel

Tab. D.3: Correlated Mundlak Chamberlain Panel for Extensive and Intensive Margins

| Variables | Correlated RE Mundlak Chamberlain Panel | | |
|---|---|---|---|
|  | LIC | MIC | Overall |
| **Intensive Margin of Log FDI** | | | |
| Log Destination Income | -0.714 | -0.400 | -0.385 |
|  | (2.47) | (0.572) | (0.517) |
| Log Source Income | 0.147 | 4.391*** | 3.874*** |
|  | (2.604) | (0.694) | (0.613) |
| Log Lending Rate | -7.018 | 0.224 | 0.308 |
|  | (4.587) | (0.533) | (0.397) |
| Log Labor productivity | 4.083 | 1.448** | 1.724*** |
|  | (3.192) | (0.705) | (0.654) |
| **Extensive Margin of FDI** | | | |
| Log Startup Costs | -0.188** | 0.0613 | -0.0281 |
|  | (0.0854) | -0.0659 | -0.0493 |
| Log Lending Rate | 0.1506 | -0.0611 | -0.0465 |
|  | (0.2205) | (0.101) | (0.093) |
| Log Labor productivity | 0.893 | 1.040*** | 0.888*** |
|  | (0.55) | (0.273) | (0.238) |
| $\hat{\sigma}$ ($\hat{\sigma_u}$/ $\hat{\sigma_\varepsilon}$) | | | |
| $\hat{\rho}$ | | | |
| Observations | 225 | | |

Note: time coverage is 2002-2010 Mundlak-Chamberlain method includes time averages of relevant explanatory variables and inverse Mills ratios from the first stage Probit regressions. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term and directional fixed effects (source and destination countries).

## D.3 Copula Function and MPK

Fig. D.1: Copula-based capital flows models

| Variables | LIC | MIC | Overall Sample |
|---|---|---|---|
|  | Copula | | |
| **Intensive Margin of FDI** | | | |
| Log Destination Income | 0.0254 | 0.599*** | 0.652*** |
|  | (0.158) | (0.0535) | (0.0477) |
| Log Source Income | 0.348*** | 0.558*** | 0.571*** |
|  | (0.0992) | (0.0486) | (0.0462) |
| Log MPK | 0.0191 | 0.802*** | -0.394* |
|  | (0.308) | (0.276) | (0.239) |
| Log Labor productivity | -0.0952 | -0.276*** | -0.0772 |
|  | (0.148) | (0.0802) | (0.130) |
| **Extensive Margin of FDI** | | | |
| Log Startup Costs | -0.119*** | -0.115*** | -0.109*** |
|  | (0.0432) | (0.0241) | (0.0248) |
| Log MPK | -0.0472 | -0.00176 | 0.134* |
|  | (0.0896) | (0.0855) | (0.0687) |
| Log Labor productivity | -0.0417 | 0.117*** | 0.156*** |
|  | (0.0378) | (0.0238) | (0.0395) |
| Log $\sigma$ | -0.206*** | 1.192*** | 1.203*** |
|  | (0.0650) | (0.0377) | (0.0424) |
| $\theta$ | -0.0198 | -2.020*** | -1.605*** |
|  | (0.0291) | (0.0909) | (0.0952) |

Note: time coverage is 2002-2010. $\theta$ refers to the dependence parameter in the copula framework. Robust standard errors in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively. All specifications include a constant term.