

Cross-border Investments, Gravity Equations and the Double-hurdle Model

Anni Norring

University of Helsinki

Faculty of Social Sciences

Economics

Master's Thesis

November 30, 2014

Tiedekunta/Osasto — Fakultet/Sektion — Faculty		Laitos — Institution — Department	
Faculty of Social Sciences		Department of Political and Economic Studies	
Tekijä — Författare — Author			
Anni Norring			
Työn nimi — Arbetets titel — Title			
Cross-border Investments, Gravity Equations and the Double-hurdle Model			
Oppiaine — Läroämne — Subject			
Economics			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's Thesis		November 2014	73
Tiivistelmä — Referat — Abstract			
<p>In this thesis I consider studying the determinants of international investments with the gravity model of international financial asset trade. I discuss the relevant literature and present a theoretical framework for gravity in cross-border investments. I compare three empirical approaches, the classic approach that studies the determinants of the observed levels of cross-border holdings by a fixed effects panel model, the dichotomous approach that studies the effects of determinants on the probability of there being a positive cross-border investment by a probit model and finally an approach which combines the two previous ones by a double-hurdle model. I propose that the double-hurdle model is the correct approach in the context of cross-border investments.</p> <p>International financial asset trade is characterized by two phenomena: tendency to overinvest in domestic assets known as home bias and a dominance of observed bilateral holdings equaling zero. The gravity model can be used to explain these phenomena by studying the determinants of cross-border investment holdings and flows. The classic approach to estimating the effect of these determinants is interested in the level of the cross-border investments observed, and although having been very successful in explaining the first phenomenon, has completely overlooked the second. An alternative and more appropriate approach is to use a model for limited dependent variables. One can treat the gravity equation as binary, and take interest in the dichotomous cross-border investment decision. However, the dichotomous approach loses valuable information by not considering the levels of the observed holdings. The double-hurdle mode is apt for the task. The double-hurdle is a model for limited dependent variables, but in addition it has a specific feature: It allows two decisions, a participation decision and a level decision, and it allows them to be determined by to separate processes. The cross-border investment decision is intuitively easy to deconstruct into two distinct decisions. An investor considering investing in a given foreign country must first make up his mind on the dichotomous question of investing or not. If, and only if, the decision is positive, he then decides on the level of his investment. There is no reason why the decision to invest should be determined by a similar process as the decision of how much to invest. Thus the double-hurdle is a correct empirical approach to estimating gravity equations of international financial asset holdings and flows.</p> <p>I illustrate my argument with an empirical application to the external assets of Finnish residents. Indeed I find that the decision to invest and the decision of how much to invest are different for Finnish investors: The classic approach to estimating the gravity equations and the dichotomous approach have qualitatively slightly different results, suggesting that the two decisions may be determined by different processes. When estimated with the double-hurdle, differences become clearer: the magnitudes of the effects, their statistical significance and in certain cases even the signs of the effects are different between the participation and the level equation. Even though these results are merely illustrative, they do point out that the double-hurdle model is well worth future research in the context of cross-border investments.</p> <p>I propose that the correct method to estimate the determinants of cross-border investments could be the double-hurdle model, previously overlooked in estimating gravity models of financial asset trade.</p>			
Avainsanat — Nyckelord — Keywords			
cross-border investments, gravity equation, double-hurdle model			

Contents

1	Introduction	1
2	Gravity model of international financial asset trade	7
2.1	Gravity in literature	7
2.2	Theoretical framework of the gravity model of cross-border investments	17
3	Estimating the determinants of cross-border investments	25
3.1	The classic approach to estimating a gravity equation	26
3.2	The dichotomous approach to estimating the gravity equation	28
3.3	Combining the two approaches: a double-hurdle model	30
4	Empirical application: the external wealth of Finland	38
4.1	Data on Finnish external wealth and its determinants	42
4.2	Methods	46
4.3	The classic approach: the determinants of the level of cross-border investments	49
4.4	The dichotomous approach: the determinants of the probability of a positive cross-border investment	54
4.5	The double-hurdle approach: the determinants of the participation and level decisions	59
5	Conclusions and proposal for future research	66
6	References	69

List of Figures

1	Total and foreign equity and debt security assets held by Finnish residents	38
2	The geographical distribution of assets held by Finnish residents in 2013	40
3	Correlation of distance and external portfolio assets held by Finnish residents	41

List of Tables

1	Share of zero holdings in observed cross-border investments of Finnish residents	39
2	Dependent variables	42
3	Independent variables	45
4	Marginal effects of the independent variables on the level of cross-border equity and debt security holdings in the classic approach	52
5	Estimation results for the level of cross-border equity and debt security holdings	53
6	Marginal effects of the explanatory factors on the probability of positive cross-border equity and debt security holdings in the dichotomous approach	57
7	Estimation results for dichotomous cross-border equity and debt security holdings	58
8	Marginal effects of the independent variables on cross-border equity holdings in the double-hurdle model	62
9	Marginal effects of the independent variables on cross-border debt security holdings in the double-hurdle model	64
10	Estimation results for the double-hurdle model of equity holdings	65

1 Introduction

In this thesis I study the determinants of cross-border financial asset holdings and flows by gravity models of international financial asset trade. In addition to discussing literature and presenting the theoretical framework of gravity in finance I compare three empirical approaches to estimating these determinants and apply the models to the case of cross-border investments of Finnish residents. I find evidence that the decisions to invest and how much to invest of Finnish residents appear to be governed by different processes. This result calls for the use of the double-hurdle model, a two-stage empirical approach to estimating determinants of cross-border investments, which has been previously unutilized in the context of cross-border investments.

The international flows and holdings of financial assets are characterized by two phenomena that are not easily explained by standard economic theory. First, home bias, or tendency to overinvest in domestic instruments, has been puzzling economists for years (see e.g. French and Poterba, 1991, for an early study and Sercu and Vanpee, 2007, for an extensive survey). Countries all over the world invest remarkably little of their financial wealth abroad and most of the economies for which data is available show a tendency to overinvest in domestic instruments. In 2005 home bias¹ varied between 30.8% for Netherlands and 99.7% for Indonesia (Sercu and Vanpee, 2007)². A small portion of the bias can be explained by the fact that external financial assets are harder to capture in official statistics than domestic assets (Zorzi, 2009, and Zucman, 2013), leading to the data on financial wealth to be biased towards domestic assets. Still, the dominant factor is likely to be that for some reason investors prefer domestic assets to foreign ones and tilt their portfolios towards domestic assets. This issue has gained a lot of interest in literature, especially after Obstfeld and Rogoff (2000) identified home

¹Here home bias in portfolio choice is defined as the difference between the observed share of domestic assets in an economy's portfolio and the share predicted by the simple CAP-model. For other definitions, see Sercu and Vanpee, 2007.

²See e.g. Li et al., 2004, and Sorensen et al., 2007, for slightly different estimates.

bias in portfolio choices to be "one of the six major puzzles in international macroeconomics". There is a lot of empirical evidence on "closeness bias", the tendency to overinvest in countries that are closer to the domestic country and the negative correlation of distance and investments, which is the standard result of the literature studying gravity in finance (see e.g. Drakos et al., 2014). Literature has also identified "friendship bias", the tendency to overinvest in countries that have tight economic and cultural relationships with the domestic country (see e.g. Berkel, 2007, and Anderson et al., 2011). There is even evidence of "local bias", home bias inside a country: Coval and Moskowitz (1999) found in their study of U.S. domestic mutual funds that investments were biased towards locally headquartered firms. Home bias thus seems to be a very fundamental feature of the financial markets.

Second, the distribution of observed flows and holdings of cross-border financial assets is skewed towards zero observations. Drakos et al. (2014) found that in the Coordinated Portfolio Investment Survey ³ 59% of all observations on bilateral cross-border investments were zero in years 2001-2007. That is, at a given period of time a given country invests in only some of the other countries of the world instead of the investments being even close to evenly distributed over possible destination countries. In addition to the tendency to understate geographical diversification in the form of home bias, investors also downplay international risk sharing by investing in only a handful of countries. Another explanation could be that there may be exogenously determined barriers to investment, which make it impossible to invest in some countries. This dominance of zero observations has not received such a rigorous attention in literature as home bias has, and instead the interest has mostly been on the positive levels of cross-border investments observed. The implicit selection mechanism, which leads to either a positive or zero cross-border asset holding, is however equally important for understanding the patterns of international investments and should also be given due attention

³CPIS-data compiled by International Monetary Fund based on the international investment position statistics of reporting countries.

(Drakos et al., 2014).

Standard economic theory struggles to explain these phenomena. Do investors simply fail to understand the benefits of geographical diversification and risk sharing? There is evidence of under-diversification being costly, reducing returns and increasing risks (see e.g. Goetzmann and Kumar, 2008). Investors are often assumed to be risk averse, so the extremely low levels of geographical risk sharing do not seem to be in accordance with theory. Could the explanation be that there are substantial barriers for investing abroad? But as asset flows have been deregulated, financial assets are not physically shipped around the world and efficient financial markets should reduce the transactions costs to minimum, the barriers to investing abroad should be negligible. What then could be the reason for such low levels of external assets?

Studying the determinants cross-border investments is of relevance if only to fill in the gaps of knowledge on external wealth. Both of the before described phenomena point to there being frictions in international financial markets which affect the cross-border investment decisions of investors. Studying the determinants of cross-border investments can give insight on these frictions. Also policy makers should be interested in the determinants of cross-border investments of investors, as there are many reasons why policy makers would want to influence the capital flows in and out of the domestic economy and the behaviour of investors. For example, if there is assumed to be a positive propensity to consume out of financial wealth, the higher returns of a better diversified portfolio could induce the domestic investors to consume more. This would have positive demand effects for the domestic economy. On the other hand, it might also be in the policy makers' interest to make the home country more appealing to domestic investors and thus, in fact, increase the degree of home bias. If, in addition, the policy makers wish to attract more foreign investments and increase capital inflow, the cross-border investment choices of investors resident in other countries should be of interest. Some of the plausible determinants of cross-border investments,

such as investment income taxation, are also instruments of financial policy, and the broad effects of policy changes should be understood. Policy makers can also influence e.g. the efficiency of transactions by tweaking the bureaucracy of repatriation of foreign returns.

The gravity model of international asset trade is currently the most widely used theoretical framework for studying the determinants of cross-border financial asset flows and holdings and understanding home bias⁴. The gravity model originates from the international goods trade literature, but after the influential studies of Martin and Rey (2004), who develop a theoretical framework for gravity model of cross-border asset flows, and Portes and Rey (2005), who estimate the model, gravity has played an important role also in studying international investment patterns.

In this thesis I discuss the literature of gravity models and present the theoretical framework of gravity in financial asset trade derived by Martin and Rey (2004). In addition, I compare three different empirical approaches to estimate the gravity model of financial asset trade. The first of these empirical models follows the work of Portes and Rey (2005) and studies the determinants of the *level* of cross-border asset trade. The second empirical model follows the work of Drakos et al. (2014), and in this set-up the interest is on the *dichotomous decision* of making a cross-border investment. The third model to be considered has to my knowledge not been employed in the context of cross-border investments before. This double-hurdle model is closely related to tobit models and other models for limited dependent variables and was first suggested by Cragg (1971) for cases where the value of dependent variable is zero with non-negligible probability. In addition to this the double-hurdle model has a specific feature: By assuming two different stages or "hurdles", a dichotomous participation decision followed by a level decision, it allows these two hurdles to be governed by different processes.

⁴For other frameworks of studying home bias see e.g. Brennan and Cao, 1997, who study the differences in informational endowments of investors, and Couerdacier et al., 2010, who develop a RBC-model explaining home bias by capital accumulation and international asset trade.

The double-hurdle model has previously been employed in set-ups that study e.g. female labor supply or consumption of cigarettes (see e.g. Blundell and Meghir, 1987, and Jones, 1989, for classic papers), both phenomenon characterized by a non-negligible share of zero observations. The context of cross-border investment decisions is a natural application for the double-hurdle model for two reasons. First, because the observations cluster on zero, i.e. the dependent variable is limited. Secondly, and more specifically, because the cross-border investment decision is intuitively easy to deconstruct into the two hurdles, participation and level decisions. An investor considering investing in a given country must first make the dichotomous decision of investing or not. Then given that the decision to invest is positive, the investor decides the level of his investment. There is no reason why these decisions should be determined by identical processes. I am indeed able to show that in the case of cross-border investment decisions of Finnish residents, the participation decision and the level decision appear to depend on different factors.

The classic approach following Portes and Rey (2005) taken by most of the studies considering the gravity model of financial asset trade fails to address the clustering of zero observations and simply ignores it. Drakos et al. (2004) address this feature of the data by using a dichotomous model, but their approach answers only half of the question. I argue that the double-hurdle model is a suitable and advisable empirical approach in the context of international asset flows as it addresses the zero observations and allows for a more realistic characterization of the investment decision. Not considering the double-hurdle model for estimating the determinants of international asset trade may result in too simplistic and possibly biased results. I illustrate this point by the case of cross-border portfolio investments of Finnish residents. Indeed, the results of the classic approach and the dichotomous approach are different, suggesting different processes are at work in determining the participation and the level decisions. The estimation results of the double-hurdle model confirm this.

I find that the determinants of the level of Finnish cross-border holdings confirm the standard results of gravity literature: The levels of holdings appear to be mainly determined by GDP of the destination economy, which has a positive effect, the bilateral distance between the economies, which has a negative effect, and some controls thought to proxy information advances and the ease of financial transactions. In contrast to this, the evidence I find in the case of the dichotomous investment decision are a lot more ambiguous. GDP of the destination country has a positive effect also on the probability of positive cross-border investments, but the effect of distance and other variables is much reduced and mixed. These qualitative differences are confirmed when the determinants of cross-border investments are estimated using the double-hurdle model. GDP of the destination economy has a positive effect throughout, but the effect of distance is clearly negative only on the level equations, and ambiguous on the participation decisions. There is thus evidence of that distance does not play such a clear role when investors decide which countries to invest in, but once this first step is settled, investors tend to invest more in countries that are nearby.

The rest of the essay is structured as follows. Section 2 discusses the literature on gravity in asset trade and presents the theoretical framework of the gravity model. Section 3 describes the empirical models and estimation methods. Section 4 presents and discusses the case of Finnish external assets, the estimations and results. Section 5 concludes.

2 Gravity model of international financial asset trade

2.1 Gravity in literature

The gravity model has been a workhorse of international trade literature for decades⁵. As the name suggests, the model, originally borrowed from physics, explains a flow between two entities by simply relating it to their two masses and a friction term. The simplest version of the model requires only the economic masses of the domestic country and the foreign country, such as GDPs, and the distance of the two countries as an approximation for transaction costs:

$$trade_{ij} = \beta_1 GDP_i + \beta_2 GDP_j + \beta_3 dist_{ij}, \quad (1)$$

where $trade_{ij}$ denotes the trade flow to country i from country j , GDP_i and GDP_j are the gross domestic products of countries i and j respectively and $dist_{ij}$ denotes the distance of the two countries. The coefficients to be estimated are β_1 , β_2 and β_3 . (Portes and Rey, 2005) In more elaborate versions the number of variables increases, but in general, in a gravity specification bilateral trade is a product of measures of economic size, a bilateral barrier (such as distance and other trade frictions) and multilateral resistance term (Anderson and van Wincoop, 2004).

Portes and Rey (1998) were one of the first to propose that a gravity equation could also explain bilateral financial asset flows. The theoretical framework for a gravity model of cross-border asset flows derived by Martin and Rey (2004) and estimated by Portes and Rey (2005) is the most influential and widely estimated approach to gravity in international finance. This micro-founded two-country model with an endogenous number of financial assets relates the size of the two economies and trade costs to their bilateral asset transactions. The model also succeeds in providing an intuitively

⁵See e.g. Anderson (1979) for an early formulation and Head and Mayer (2013) and Anderson (2010) for recent surveys.

appealing explanation for home bias by assuming that transaction costs enter the model non-linearly, enabling even small transaction costs to produce a high degree of home bias. A theory of financial asset trade which gives rise to a gravity equation able to produce high degrees of home bias has the following key elements:

1. Assets are imperfect substitutes in the sense that they provide insurance against different types of risks.
2. International asset trade entails transaction costs.
3. Supply of assets is endogenous⁶.
4. There is imperfect competition on the asset markets.
5. Transaction costs interact non-linearly with the elasticity of substitution between assets.

In the model risk-averse agents develop risky projects of Arrow-Debreu nature (see e.g. Obstfeld and Rogoff, 1999), i.e. that have a positive payoff in only one state of nature. Shares of the projects are then traded on the stock market. Thus the market capitalization in both countries is endogenous. Higher aggregate demand from the foreign country induces higher asset prices in the domestic market, thus increasing the optimal amount of risky projects and financial assets. Their model gives rise to a specification of the total value of bilateral asset flows including transaction costs. After taking logs and using the equilibrium demand for foreign assets from the model this specification takes the form of a gravity equation, which relates the flow of financial assets to economic masses of the two countries and friction terms:

$$\log T_{ij} = \log(p_j n_j z_j) + \log(n_i c_{1i}) + \log \phi + \sigma \log\left(\frac{d}{L p_j}\right) + \log\left(\frac{\beta^\sigma}{d}\right). \quad (2)$$

T_{ij} denotes the investment flow from country i to country j . The first term on the right-hand side refers to the financial wealth of the foreign economy

⁶See e.g. Coeurdacier and Martin (2009) for a simplification of the model, where the supply of assets is assumed exogenous.

and the second to aggregate consumption in the domestic economy in the first period. The third term captures the effect of transactions costs, the fourth that of expected returns from the foreign assets. The last term is a constant. (Martin and Rey, 2004)

In addition to the theoretical framework of Martin and Rey (2004), there are two alternative ways to derive a theoretical equation of cross-border asset trade discussed in the literature. One is proposed by Lane and Milesi-Ferretti (2008). Their model is an N-country extension of a model by Obstfeld and Rogoff (2001) in which bilateral trade flows are linked to bilateral asset holdings. In this framework the barriers to goods trade result in home bias in financial assets. This model however has not been empirically confirmed and indeed seems to be plausible only in theory, because the existence of a real exchange rate hedge channel, crucial to the model, which allows barriers in goods trade to affect asset trade, has not been found to be working in practice (Okawa and van Wincoop, 2012).

Another theoretical framework is proposed by Okawa and van Wincoop (2012), who establish theoretical foundations of a gravity model analogous to the gravity model of international goods trade. The authors point out the limitations of the approach of Martin and Rey (2004) and consider the majority of papers estimating the determinants of cross-border asset flows to be on a shallow theoretical ground. They criticize the unrealistic assumption of Arrow-Debreu-assets, as most financial assets have a positive payoff in multiple states of nature, and the inclusion of expected return in the gravity equations. They develop a gravity model for bilateral asset holdings in a one-good, two-period and N-country framework and show that by introducing fairly simple extensions to the model, the gravity form is lost. At the moment there are little empirical studies to back up this model, while the model of Martin and Rey (2004) is tested in a multitude of studies. It thus seems sensible to for the present take the micro-foundations for the gravity model as derived by Martin and Rey (2004), bearing in mind the limitations of their approach.

Portes and Rey (2005) estimate the model of Martin and Rey (2004) and find that the model indeed fits data remarkably well. Financial asset flows seem to be mainly determined by market sizes of economies and trading costs, of which information costs make up a substantial share. These information costs seem to be well approximated by simply the distance between the two countries, but the authors also proxy information frictions by other more explicitly information related variables: telephone traffic, number of subsidiaries of source country's banks in the destination countries, extent of insider trading in the destination country and common trading hours. Besides information costs, transaction costs are also a function of the efficiency of transaction technology. To proxy this effect, the authors use an index of financial market sophistication of both countries. Their dependent variable is the level of equity of destination country held by residents of the source country. Thus, the gravity equation to be estimated takes the following form:

$$\begin{aligned}
\log T_{ij,t} = & \alpha_1 \log(mktcap_{i,t}) + \alpha_2 \log(mktcap_{j,t}) + \alpha_3 \log(dist_{ij}) \\
& + \alpha_4(\text{information variables}) \\
& + \alpha_5(\text{transaction technology variables}) \\
& + \text{time dummies} + \text{constant} + \epsilon_{ij,t}, \tag{3}
\end{aligned}$$

where $T_{ij,t}$ refers to the bilateral transactions in equity from source country to destination country at t , $mktcap$ to equity market capitalization and $dist_{ij}$ to the distance between the capital of the two countries. $\epsilon_{ij,t}$ is the error term. Theory predicts that the coefficients of the logs of economic masses are equal to one, and that distance has a negative effect on asset transactions, while information and transaction technology have a positive effect. This means that $\alpha_1 = \alpha_2 = 1$, $\alpha_3 < 0$, $\alpha_4 > 0$ and $\alpha_5 > 0$, and the empirical equation can be estimated as:

$$\begin{aligned}
\log\left(\frac{T_{ij,t}}{mktcap_{i,t} * mktcap_{j,t}}\right) = & \beta_1 \log(dist_{ij}) + \beta_2(\text{information variables}) \\
& + \beta_3(\text{transaction technology variables}) \\
& + \text{time dummies} + \text{constant} + \nu_{ij,t}, \tag{4}
\end{aligned}$$

where $\beta_1 < 0$, $\beta_2 > 0$, $\beta_3 > 0$ and $\nu_{ij,t}$ is the error term. This specification is arguably simple, but it has been very successful in explaining the determinants of cross-border asset flows with economic relevance attached to the estimated coefficients. This has led to the contribution of Portes and Rey (2005) being instrumental in the "explosion of papers estimating gravity equations for cross-border financial holdings" (Okawa and van Wincoop, 2012), which has also been encouraged by the extensive CPIS-dataset of IMF published from 2001 onwards.⁷

Portes and Rey (2005) elaborate their model by studying the effects of bilateral trade on cross-border investment flows, but their results rule out financial asset flows being simply a mirror image of international trade contrary to the traditional view. Martin and Rey (2004) or Portes and Rey (2005) do not consider the multilateral resistance term interpreted as the effect of all competing assets on the demand of the foreign asset and emphasized in international goods trade literature (see e.g. Anderson and van Wincoop, 2004) and this is one of the weaknesses of the Martin-Rey-model pointed out by Okawa and van Wincoop (2012). E.g. Couerdacier and Martin (2009) extend their model by deriving multilateral resistance term in the theoretical framework and including it in the empirical application in the investor country fixed effects. Drakos et al. (2014) address the multilateral resistance term by simply including regional dummies and a full set of country-pair-dummies. Portes and Rey (2005) do include full set of country-dummies in one of their specifications, but they do not motivate this by the multilateral resistance term, but by addressing unobserved country-specific features.

The theoretical framework of Martin and Rey (2004) and its empirical application of Portes and Rey (2005) study *flows* of cross-border investments, but the framework is fully applicable to studying *stocks* of assets as well. In fact, as the CPIS-dataset is the main source of data for many studies on gravity models and this data includes only stocks of cross-border portfolio

⁷See also Coeurdacier and Guibaud, 2011, for a convenient categorization of different studies.

assets, most empirical studies of the gravity model consider stocks of external assets (e.g. Li et al., 2004, Aviat and Coeurdacier, 2007, Coeurdacier and Martin, 2009, Balli et al., 2013, among others). However, with stock of holdings the problem of serial persistence in the cross-border asset holdings becomes even more acute. Drakos et al. (2014) point out that the feature of serial correlation is utterly overlooked in previous literature, even though it is clear that cross-border asset holdings are strongly dependent on their previous values. The most appropriate approach would thus be to model the cross-border asset holdings using dynamic models, although this is beyond the scope of this thesis.

There have also been different specifications for the models stemming from the theoretical framework of Martin and Rey (2004) proposed in the literature. Couerdacier and Martin (2009) extend the model of Martin and Rey (2005) to N countries and show that assuming exogenous asset supply instead of endogenous has little effect on the results. In their model the authors include a multilateral resistance term, which can be thought to be a measure of financial remoteness. The term takes the form of a price index for all the assets that compete with the imported asset. Their theoretical gravity equation for $Asset_{ij}$, the financial assets of country j held by country i , takes the form:

$$Asset_{ij} = L_i p_j n_j s_{ij} = \frac{\beta L_i y_i n_j}{1 + \beta} \left(\frac{r_j Q_i}{\tau_{ij}} \right)^{\sigma-1} \quad (5)$$

where $L_i y_i$ refers to GDP of the source country, n_j to market capitalization of the host country, τ_{ij} to the transaction costs, r_j to expected return of the foreign asset, σ to the elasticity of substitution between assets and Q_i is the multilateral resistance term of the form:

$$Q_i = \left[\sum_{j=1}^N n_h \left(\frac{r_j}{\tau_{ij}} \right)^{\sigma-1} \right]^{\frac{1}{1-\sigma}}. \quad (6)$$

By taking logs equation (5) takes the familiar gravity form:

$$\log(Asset_{ij}) = \log L_i y_i + \log n_j - (\sigma - 1) \log \tau_{ij} + (\sigma - 1) \log r_j + (\sigma - 1) \log Q_i, \quad (7)$$

which Couerdacier and Martin (2009) test by estimating the following equation:

$$\log(Asset_{ij}) = \alpha_i + \beta \log(GDP_j) + \gamma \frac{mktcap_j}{GDP_j} + (\sigma - 1) \log Z_{ij} + (\sigma - 1) \log r_j \quad (8)$$

where α_i is the investor country fixed-effect and Z_{ij} denotes the transaction costs on the international financial markets. GDP of the investor country and the multilateral resistance term Q_i are controlled for in the investor country fixed-effect α_i . Coeurdacier and Martin (2009) apply their specification of the gravity model to studying the impact of euro on the determinants of trade in bonds, equity and banking assets.

The approach taken by Aviat and Couerdacier (2007) combines the gravity model of bilateral goods trade and bilateral financial asset trade. They derive a system of gravity equations, where trade in goods explains bilateral financial asset holdings and asset holdings explain bilateral goods trade flows. The following two equations specify a system where cross-border asset holdings and international trade in goods are mutually determined:

$$\begin{aligned} \log \frac{Asset_{ij}}{GDP_i * GDP_j} = & \alpha_A + \phi_A \log \frac{Trade_{ij}}{GDP_i * GDP_j} - \beta_A \log dist_{ij} \\ & + \gamma_A Z_{ij,A} + \sigma_A r_j + \epsilon_{ij,A} \end{aligned} \quad (9)$$

$$\begin{aligned} \log \frac{Trade_{ij}}{GDP_i * GDP_j} = & \alpha_T + \phi_T \log \frac{Asset_{ij}}{GDP_i * GDP_j} - \beta_T \log dist_{ij} \\ & + \gamma_T Z_{ij,T} + \epsilon_{ij,T} \end{aligned} \quad (10)$$

where $Asset_{ij}$ refers to bilateral asset holdings, $Trade_{ij}$ to bilateral goods trade, α to fixed effects in asset trade (A) and in goods trade (T), $dist_{ij}$ to bilateral distance, Z_{ij} to trade costs, r_j to returns on foreign assets and ϵ_{ij} is error term. The authors find that controlling for trade, the impact of distance on assets holdings is drastically reduced. This is very different from Portes

and Rey (2005) who find that trade had very little effect on their results. Also Balli et al. (2013) find that bilateral trade flows, measured as the sum of exports and imports, are the strongest determinants of bilateral asset holdings for all sectors, including households. But why should trade flows have an effect on the investment decisions of households? To my understanding, much of the significance of trade flows for cross-border investment holdings might actually also be related to "familiarity effects" induced by trade flows, instead of trade flows *per se*.

The gravity model may be disaggregated over investor sectors, but to my knowledge only one study so far considers the cross-border investments of different sectors. Balli et al. (2013) apply the gravity model on the cross-border assets of financial, insurance, non-financial, government and household sectors instead of the aggregate economy. The CPIS data allows this kind of study for all countries, which report the external assets disaggregated over sectors. Examining the determinants of international asset flows separately by sectors makes sense, because it is likely that the determinants of cross-border decisions, or at least their coefficients, of e.g. sophisticated institutional investors are different from those of households (see e.g. Grinblatt and Keloharju, 2000). Balli et al. (2013) indeed find this to be the case.

An interesting and original empirical approach is taken by a very recent paper by Drakos et al. (2014). Based on the gravity set-up of Martin and Rey (2004), the authors employ random effects probit model to explain the substantial number of zero observations in global bilateral investment holdings. Instead of concentrating on the level of bilateral asset holdings, the authors are interested in the dichotomous decision of investing or not investing. The authors argue that this way they are able to make use of the data in a much fuller sense, as now also the confidential data and zero holdings yield important information. They estimate the following empirical model:

$$Inv_{i,t} = 1\{\gamma Inv_{i,t-1} + \mathbf{X}'_{i,t}\boldsymbol{\beta} + u_{i,t} > 0\}, \quad (11)$$

$$i = 1, \dots, N; t = 1, \dots, T,$$

where indicator function $1\{\}$ takes the value one when the event inside the brackets occurs and zero otherwise. $Inv_{i,t}$ denotes the dichotomous dependent variable. This set-up has a cross-sectional dimension of pairs of source-host countries i , and a time dimension. The vector $\mathbf{X}_{i,t}$ includes covariates which model the discrete choice of investing in a foreign financial market: host country specific, $x_{i,h,t}$, source country specific, $x_{i,s,t}$, and pair specific, x_i , characteristics. These variables include the economic masses of the countries and transaction costs. Economic masses are measured by GDP per capita. The role of transaction costs is taken by information effects and attractiveness to investments. Information frictions are proxied by distance and information advantages by common language and common legal origins. Attractiveness to investments is measured by openness of the destination country's economy, its institutional quality approximated by indices of financial and political risk and financial market development measured by stock turnover ratio and domestic credit per GDP. Their empirical specification also includes the multilateral resistance term, which is interpreted as a price index of all financial assets which are (incomplete or complete) substitutes for the foreign asset. The term is estimated by regional dummies for continents and a full set of source and host-specific dummies. The authors estimate a static and a dynamic version of the model. The static version of the model is estimated by random effects probit model and it is assumed that there is no state dependence and no serial correlation in the transitory errors. The static model thus concentrates solely on the roles of the independent variables in determining the probability of a positive investment from the source country to the destination country. The standard result of gravity models is confirmed in the discrete choice framework: distance and positive bilateral cross-border investments are negatively linked. There is thus a smaller probability of investing in a distant destination country, all things being equal. (Drakos et al., 2004)

I take the work of Martin and Rey (2004) as the theoretical base of my study because of the merits discussed earlier: The model has been extensively

tested and found to fit the data of financial asset flows reasonably well. The empirical approaches to estimating gravity equations to be considered further are that of Portes and Rey (2005), because their approach can duly be considered the classic approach to gravity in finance, and Drakos et al. (2004), because of their original contribution to the study of cross-border investment. Importantly, these two models are chosen also because they can be thought to represent the two hurdles of the double-hurdle model. According to Cragg (1971) the participation decision can be represented by a probit model and the level decision by a fixed effects model. Thus the dichotomous approach of Drakos et al. (2014) corresponds to the participation decision and the classic approach to the level decision.

2.2 Theoretical framework of the gravity model of cross-border investments

The model of Martin and Rey (2004) is a two-period model with two economies; domestic, i , and foreign, j . The economies are populated by n_i and n_j risk-averse immobile agents. In the first period the agents are endowed with y_i and y_j of tradable good, which can be consumed or saved. Saving occurs through investing in projects either directly by investing to developing a project or indirectly by buying shares of the projects on the stock market. Each agent develops z_i and z_j projects in the domestic and foreign economy respectively, and the set of projects developed by domestic and foreign agents are denoted respectively by Z_i and Z_j . Agents in both economies are identical and thus develop the same number of projects. The aggregate amount of projects developed depends on the population size and thus the number of projects developed in the domestic and the foreign economy respectively are $n_i z_i$ and $n_j z_j$. The international equilibrium total number of assets is $n_i z_i + n_j z_j$. The sets of domestic and foreign projects are denoted respectively by M_i and M_j . The total costs of developing a domestic and foreign project are $f(z_i)$ and $f(z_j)$ respectively. The marginal costs of developing new projects are assumed to be growing because of e.g. monitoring costs that grow as the number of projects grow.

In the second period there are L exogenously determined and equally likely states of nature, which affect the dividends paid by projects in the following manner:

$$\text{Project } l \text{ pays } \begin{cases} d & \text{if state} = l \in \{1, \dots, L\}, \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

This structure of payoffs makes investing in a specific project similar to buying an Arrow-Debreu-asset (see e.g. Obstfeld and Rogoff, 1991). Arrow-Debreu-instruments have a positive payoff in only one state of nature, and it is also assumed that in every state of nature only one instrument has a

positive payoff⁸. Dividends are the only source of income in the second period. The shares of different projects are imperfectly correlated, and there is thus an incentive to diversify. Importantly, it is assumed that markets are incomplete:

$$n_i z_i + n_j z_j < L. \quad (13)$$

That is, the total number of projects is less than the number of states of nature. It is thus not possible to eliminate all risk by holding a portfolio that includes all assets. For this reason there are some states of nature, where there are no dividends, thus no income and zero consumption.

All agents are identical and the projects they develop are symmetrical. Thus demands for assets in a given economy are symmetric, and the prices and diversification choices of the agents identical. It is thus possible to omit reference to an individual agent in the notation, and only refer to the nationality of the agent.

Transaction costs When agents trade assets internationally, they face a transaction cost τ . This cost is assumed to be a so called iceberg⁹ cost, familiar from goods trade: it is paid in units of the asset itself and in proportion to the asset. The transaction cost is beared by the investors: when a domestic investor buys a foreign asset, the total cost of the investment is $(1 + \tau)p_j s_{ij}$, where p_j is the price of the foreign asset and s_{ij} is the amount of foreign assets demanded by a domestic investor. The transaction costs are understood in the broadest sense to include e.g. banking commissions, exchange-rate costs and importantly also information costs.

⁸Okawa and van Wincoop (2012) see this assumption as the main limitation of the approach of Martin and Rey, because most financial assets, such as equities and bonds, have non-zero payoffs in multiple states of nature. Moreover, in contrast to an Arrow-Debreu-asset, different financial assets have non-zero payoffs in similar states of nature. This limitation is understood and acknowledged.

⁹The analogy to iceberg is that part of the asset can be thought to "melt" during transaction (Martin and Rey, 2004).

Budget constraint The budget constraint of the first period for an agent in the domestic country is

$$c_{1i} + f(z_i) + \sum_{\substack{i \in M_i \\ i \notin Z_i}} p_i s_{ii} + \sum_{j \in M_j} (1 + \tau) p_j s_{ij} = y_i + \sum_{k \in Z_i} p_{ik} \alpha_{ik}. \quad (14)$$

The first term on the left-hand side refers to consumption in the first period and the second, third and fourth term refer to investments made through developing projects, buying domestic assets and buying foreign assets. On the right-hand side y_i refers to agents endowment of a tradable good, and the second term to income from selling shares of developed projects on the stock market. Coefficient α_{ik} is the share of the projects developed by the agent himself that he chooses to sell and is interpreted as a diversification factor.

The budget constraint of the second period depends on the state of nature:

$$c_{2i} = \begin{cases} ds_{i,l} & \text{if state} = l \in M_i, \notin Z_i, \\ (1 + \tau) ds_{i,l} & \text{if state} = l \in M_j, \\ d(1 - \alpha_{i,l}) & \text{if state} = l \in Z_i, \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

In the first case a domestic project developed by another agent succeeds and pays dividends, in the second a foreign project and in the third a project developed by the agent himself. In the fourth case the agent did not invest in a succesful project and does not receive any income in the second period.

The budget constraints are symmetrical for an agent from the foreign economy with only i replaced with j .

Utility function Expected utility of a domestic agent is of CRRA-form¹⁰:

$$EU_i = \frac{c_{1i}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \beta E \left(\frac{c_{2i}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} \right), \quad (16)$$

¹⁰Constant relative risk aversion utility function, see e.g. Obstfeld and Rogoff, p. 278-279.

where β denotes discount rate, $\frac{1}{\sigma}$ the degree of risk aversion and σ the elasticity of substitution between assets. Given the different states of nature and the assumption that these all have the same probability $\frac{1}{L}$, the expected utility takes the form

$$\begin{aligned}
EU_i = & \frac{c_{1i}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \frac{D}{1-\frac{1}{\sigma}} \sum_{\substack{i \in M_i \\ i \notin Z_i}} s_{ii}^{1-\frac{1}{\sigma}} + \frac{D(1+\tau)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} \sum_{j \in M_j} s_{ij}^{1-\frac{1}{\sigma}} \\
& + \frac{D}{1-\frac{1}{\sigma}} \sum_{k \in Z_i} (1-\alpha_{ik})^{1-\frac{1}{\sigma}} + 0, \tag{17}
\end{aligned}$$

where $D \equiv \beta d^{1-\frac{1}{\sigma}}/L$.

Maximization problem Agents maximize expected utility (17) given the intertemporal budget constraint given by (14) and (15) by choosing consumption in the first period, c_{1i} , number of projects individually developed, z_i , demand for domestic and foreign assets, s_{ii} and s_{ij} , and the degree of diversification, α_{ik} :

$$\begin{aligned}
& \max. EU_i \text{ s.t. intertemporal budget constraint} \\
& \text{w.r.t. } c_{1i}, z_i, s_{ii}, s_{ij} \text{ and } \alpha_{ik}. \tag{18}
\end{aligned}$$

Asset demands The individual asset demands of domestic agents are given by the first order conditions of the maximization problem:

$$\text{Demand for domestic assets: } s_{ii} = c_{1i} p_i^{-\sigma} D^\sigma, \quad i \in M_i, i \notin Z_i \tag{19}$$

$$\text{Demand for foreign assets: } s_{ij} = c_{1i} p_j^{-\sigma} D^\sigma \frac{(1-\tau)^{\sigma-1}}{(1+\tau)^\sigma}, \quad j \in M_j. \tag{20}$$

Note that demand for foreign assets is decreasing in transaction costs in a non-linear way. This is an important feature of the model, because it allows for high degree of home bias with relatively small transaction costs.

Market structure and asset supplies Projects have growing marginal costs and in addition there is a fixed cost related to developing a new project. Due to imperfect competition on the asset markets, a new project gives its developer monopoly power. It is thus always more profitable to develop a new project and never optimal to replicate an existing project. The supply of shares is characterized by monopolistic competition: the number of projects developed by each agent is small relative to the total number of projects. The individual supplies of shares given by the first order conditions are:

$$\text{Supply of domestic assets: } \sum_{k \in Z_i} \alpha_{ik} = \alpha_i = 1 - c_{1i} \delta p_i^{-\sigma} D^\sigma \quad (21)$$

$$\text{Supply of foreign assets: } \sum_{k \in Z_j} \alpha_{jk} = \alpha_j = 1 - c_{1j} \delta p_j^{-\sigma} D^\sigma, \quad (22)$$

where $\delta \equiv \left(\frac{\sigma}{\sigma-1}\right)^\sigma \geq 1$. Coefficient δ can be interpreted as a measure of imperfect competition, which takes value 1 when there is perfect competition.

Equilibrium conditions In equilibrium the amount of shares offered, α_i and α_j , equal the sum of aggregate domestic demand, $n_i s_{ii}$, and aggregate foreign demand, $n_i s_{ij}$, including transaction costs. The equilibrium conditions for each stock market thus are:

$$\text{Domestic stock market: } \alpha_i = (n_i - 1) s_{ii} + (1 + \tau) n_j s_{ji} \quad (23)$$

$$\text{Foreign stock market: } \alpha_j = (n_j - 1) s_{jj} + (1 + \tau) n_i s_{ij}, \quad (24)$$

The levels of shares listed on the stock market can be derived from these equilibrium conditions (23) and (24) and the first order conditions as:

$$\text{Share of domestic projects listed: } \alpha_i = \frac{(n_i - 1) c_{1i} + n_j \phi c_{1j}}{(n_i - 1 + \delta) c_{1i} + n_j \phi c_{1j}} \quad (25)$$

$$\text{Share of foreign projects listed: } \alpha_j = \frac{(n_j - 1) c_{1j} + n_i \phi c_{1i}}{(n_j - 1 + \delta) c_{1j} + n_i \phi c_{1i}}, \quad (26)$$

where $\phi = \left(\frac{1-\tau}{1+\tau}\right)^{\sigma-1}$. Coefficient ϕ is interpreted as an indicator of market segmentation, as it measures the extent to which the transaction cost and elasticity of substitution between assets lead to diversification. Lower τ implies lower segmentation, indicated by a higher ϕ .

Asset prices Asset prices are solved from the equilibrium conditions (23) and (24) and the first order conditions:

$$\text{Price of domestic assets: } p_i = d[(n_i - 1 + \delta)c_{1i} + n_j\phi c_{1j}]^{\frac{1}{\sigma}} \quad (27)$$

$$\text{Price of foreign assets: } p_j = d[(n_j - 1 + \delta)c_{1j} + n_i\phi c_{1i}]^{\frac{1}{\sigma}}. \quad (28)$$

Number of assets The number of projects is optimal when the marginal cost of the last project equals the price of the asset. Thus the optimal z_i and z_j are given by $f'(z_i) = p_i$ and $f'(z_j) = p_j$.

Size and income effects The respective sizes of the economies have effect on the proportion of each project sold on the market and the asset prices. From (25) and (26) it can be seen that if $n_i > n_j$, then $\alpha_i > \alpha_j$. That is, in a bigger economy, project owners sell a larger share of their projects on the stock market. This can be interpreted as financial market development depending on market size. From (27) and (28) it can be seen that if $n_i > n_j$, then $p_i > p_j$, so a bigger market size also results in higher asset prices. As the cost function is convex, the larger economy with higher asset prices can also sustain more projects per agent: $z_i > z_j$. This leads to there being a broader supply of different assets on the larger market.

In addition to size effects, there are also income effects in the model. From (27) and (28) it can be seen that an increase in domestic income and thus consumption increases the price of domestic assets, and from (25) and (26) that it decreases the share of projects sold on the stock market.

Demand and supply curves of assets Demand curves are solved from the equilibrium conditions (23) and (24) and individual asset demands (19) and (20):

$$\alpha_i = p_i^{-\sigma} D^\sigma [(n_i - 1)c_{1i} + n_j \phi c_{1j}] \quad (29)$$

$$\alpha_j = p_j^{-\sigma} D^\sigma [(n_j - 1)c_{1j} + n_i \phi c_{1i}] \quad (30)$$

Supply curves are given by the individual supplies (21) and (22):

$$\alpha_i = 1 - c_{1i} \delta p_i^{-\sigma} D^\sigma \quad (31)$$

$$\alpha_j = 1 - c_{1j} \delta p_j^{-\sigma} D^\sigma \quad (32)$$

Home bias By definition, there is a positive degree of home bias when the share of domestic assets in an economy's portfolio is larger than the share of domestic assets of all assets in the world. There are also other definitions of home bias, but this is the most simple and general of the definitions (see e.g. Sercu and Vanpee, 2007). In terms of the gravity model this can be formulated as the difference between the share of domestic assets of all assets held by domestic agents and the share of all domestic assets from the sum of domestic and foreign assets. Thus, there is home bias, when

$$\frac{p_i z_i (n_i - 1) s_{ii} + p_i z_i (1 - \alpha_i)}{p_i z_i (n_i - 1) s_{ii} + p_i z_i (1 - \alpha_i) + (1 + \tau) p_j z_j n_j s_{ij}} > \frac{n_i p_i z_i}{n_i p_i z_i + n_j p_j z_j}$$

$$\Leftrightarrow n_i n_j (1 - \phi^2) + (n_i + n_j) (\delta - 1) + (\delta - 1)^2 > 0$$

$$\Leftrightarrow \text{there is home bias when } \phi \neq 1 \text{ or } \delta \neq 1, \quad (33)$$

where $\phi = \left(\frac{1-\tau}{1+\tau}\right)^{\sigma-1}$ and $\delta = \left(\frac{\sigma}{\sigma-1}\right)^\sigma$. Coefficient ϕ is interpreted as an indicator of market segmentation and δ as a measure of imperfect competition.

The non-linear interaction of transaction costs and elasticity of substitution between assets gives an intuitively appealing explanation to home bias, and the high levels of home bias observed can be caused by low transactions costs when modeled this way.

Financial asset trade flows and the gravity equation Finally, the model gives rise to an equation of bilateral asset flows, which takes a form similar to gravity equation of goods trade. This gravity equation of bilateral financial asset flows provides a theoretical framework for analyzing the determinants of cross-border investment decisions.

The total value of investments of domestic agents in foreign assets including transaction costs is given by

$$T_{ij} = n_i s_{ij} p_j n_j z_j (1 + \tau), \quad (34)$$

where i refers to domestic economy, j to foreign economy, T_{ij} to assets of j held by residents of i , n_i and n_j to populations of the economies, p_j to foreign assets price, $n_j z_j$ to the number of assets in the foreign economy, s_{ij} to demand for foreign assets by residents of the domestic economy and τ to transaction costs. By taking logs and using (20), the equilibrium demand for foreign assets, the equation takes a gravity form familiar from international goods trade:

$$\log T_{ij} = \log(p_j n_j z_j) + \log(n_i c_{1i}) + \log \phi + \sigma \log\left(\frac{d}{L p_j}\right) + \log\left(\frac{\beta^\sigma}{d}\right), \quad (35)$$

where $\phi = \left(\frac{1-\tau}{1+\tau}\right)^{\sigma-1}$ and σ is the elasticity of substitution between assets. The first term on the right-hand side refers to the financial wealth of the foreign economy and the second to aggregate consumption in the domestic economy in the first period. The third term captures the effect of transactions costs through the transformation of transaction costs, ϕ . The fourth term is a function of expected returns of the foreign assets, where d denotes dividends determined by (12). The last term is a constant, where β denotes discount rate and d payoff.

Martin and Rey (2004) interpret this equation for flows of investments. Interpreting it for stocks of financial assets held by domestic investors requires no changes in the formulation. It should be noted that none of these terms can be readily interpreted as the multilateral resistance term, which is included in the specification of Couerdacier and Martin (2009) and emphasized by trade literature (see e.g. Okawa and van Wincoop, 2012).

3 Estimating the determinants of cross-border investments

As discussed in the literature review in section 2.1 there are a number of different empirical approaches to estimating the determinants of a gravity equation of financial asset trade that have their theoretical origins in the framework of Martin and Rey (2004). I have chosen two of these for a closer look. The empirical model of Portes and Rey (2005) has formed the basis of the vast majority of models studying the gravity framework, and I refer to it as the "classic approach". A very different empirical approach is taken in a recent paper by Drakos et al (2014), referred to here as the "dichotomous approach". Instead of concentrating on the level of bilateral asset holdings, the authors are interested in the binary decision of investing or not investing. In addition to presenting these two empirical models, I will propose a third one, the double-hurdle model, which has not been employed before in the context of international asset flows.

A simplified version of (35), the gravity equation of Martin and Rey (2004), is used as a theoretical framework and a starting point by e.g. Portes and Rey (2005) and also the empirical part of this thesis. This version of the gravity equation states the log of transactions in financial assets T_{ij} from country i to country j as

$$\begin{aligned} \log T_{ij,t} &= k_1 \log(M_{i,t} M_{j,t}) + k_2 \log(\tau_{ij,t}) + k_3, \\ i, j &= 1, \dots, N \text{ and } t = 1, \dots, T. \end{aligned} \tag{36}$$

$M_{i,t}$ and $M_{j,t}$ are measures of the economic masses of country i and j , while $\tau_{ij,t}$ denotes transaction costs between the countries. Coefficients k_1 and k_2 are to be estimated and k_3 a constant. There is a cross-sectional dimension of pairs of source-host countries ij and a time dimension t .

3.1 The classic approach to estimating a gravity equation

The empirical approach of e.g. Portes and Rey (2005), referred to here as the classical approach, takes the simplified theoretical gravity equation (36) as a starting point and relates the level of foreign assets held by domestic investors to the economic masses of the two countries and broadly understood transaction cost variables:

$$\begin{aligned} \log T_{ij,t} &= \alpha_1 \log M_{i,t} + \alpha_2 \log M_{j,t} + \alpha_3 \log \tau_{ij,t} + u_{ij,t}, \\ i, j &= 1, \dots, N \text{ and } t = 1, \dots, T. \end{aligned} \quad (37)$$

$T_{ij,t}$ refers to the investments from source country i to destination country j , $M_{i,t}$ and $M_{j,t}$ to economic masses of the two countries and $\tau_{ij,t}$ to the transaction costs between the two countries. $u_{ij,t}$ is the error term. There is a cross-sectional dimension of pairs of source-host countries ij and a time dimension t . The effect of transaction costs is approximated by distance as a proxy of information frictions, information variables and transaction technology variables:

$$\begin{aligned} \log T_{ij,t} &= \alpha_1 \log(M_{i,t}) + \alpha_2 \log(M_{j,t}) + \alpha_3 \log(\text{dist}_{ij,t}) \\ &+ \alpha_4 \log(\text{information variables}) \\ &+ \alpha_5 \log(\text{transaction technology variables}) \\ &+ \text{time dummies} + \text{country dummies} + \text{constant} + u_{ij,t}, \\ i, j &= 1, \dots, N \text{ and } t = 1, \dots, T. \end{aligned} \quad (38)$$

$T_{ij,t}$, $M_{i,t}$ and $M_{j,t}$ are as before and $\text{dist}_{ij,t}$ refers to the distance between the capitals of the two countries. Theory predicts that the coefficients of the logs of economic masses are equal to one, and that distance has a negative effect on asset transactions, while information and transaction variables have a positive effect. Thus it is expected that $\alpha_1 = 1$, $\alpha_2 = 1$, $\alpha_3 < 0$, $\alpha_4 > 0$ and $\alpha_5 > 0$. Time dummies capture the effect of macroeconomic disturbances and country dummies the effect of the multilateral resistance term (Coeurdacier and Martin, 2009).

The coefficients α_1 , α_2 , α_3 , α_4 and α_5 of this empirical gravity equation (38) are estimated through fixed effects estimation for panel data. As all variables are in logs, the resulting coefficients can be interpreted as elasticities, that is, ratios of the percentage change in the dependent variable to the percentage change in independent variables. (Portes and Rey, 2005)

The dependent variable is assumed not to be autocorrelated, but it is clear that this is not a realistic assumption at least in the case of asset holdings (see e.g. Drakos et al., 2014). Including lagged values of the dependent variable would require a dynamic panel data model, as including them in fixed effects estimation would result in inconsistent estimates. The most appropriate way would be to employ a dynamic panel data model, but this is alas beyond the scope of this thesis.

3.2 The dichotomous approach to estimating the gravity equation

Drakos et al. (2014) take a different empirical approach to estimating the gravity equation of the theoretical set-up of Martin and Rey (2004). In this approach, referred to here as the dichotomous approach, the interest is in the dichotomous decision of investing or not. With the simplified theoretical gravity equation (36) again as a starting point an empirical dichotomous gravity equation is formulated as a binary choice panel data model:

$$\begin{aligned}
 DT_{ij,t} &= 1\{T_{ij,t} > 0\} \\
 &= 1\{\alpha_1 \log M_{i,t} + \alpha_2 \log M_{j,t} + \alpha_3 \log \tau_{ij,t} + u_{ij,t} > 0\}, \quad (39) \\
 &i, j = 1, \dots, N \text{ and } t = 1, \dots, T.
 \end{aligned}$$

$DT_{ij,t}$ is the dichotomous dependent variable, $T_{ij,t}$ refers to the investments from source country i to destination country j , $M_{i,t}$ and $M_{j,t}$ to economic masses of the two countries and $\tau_{ij,t}$ to the broadly understood transaction cost variables between the two countries. $u_{ij,t}$ is the error term. The indicator function $1\{\}$ takes the value one when the event inside the brackets occurs, that is when the observed cross-border investments are positive, and zero otherwise. There is a cross-sectional dimension of pairs of source-host countries ij and a time dimension t .

As in the classic approach, the effect of transaction costs is approximated by distance as a proxy of information frictions, and information variables and transaction technology variables:

$$\begin{aligned}
 DT_{ij,t} &= 1\left\{\alpha_1 \log(M_{i,t}) + \alpha_2 \log(M_{j,t}) + \alpha_3 \log(dist_{ij}) \right. \\
 &\quad + \alpha_4 \log(\text{information variables}) \\
 &\quad + \alpha_5 \log(\text{transaction technology variables}) \\
 &\quad + \text{time dummies} + \text{country dummies} \\
 &\quad \left. + \text{constant} + u_{ij,t} > 0\right\}, \quad (40) \\
 &i, j = 1, \dots, N \text{ and } t = 1, \dots, T.
 \end{aligned}$$

$DT_{ij,t}$, $T_{ij,t}$, M_i , M_j and $dist_{ij}$ are as before. Time dummies capture the effect of macroeconomic disturbances and country dummies the effect of the multilateral resistance term. The coefficients α_1 , α_2 , α_3 , α_4 and α_5 of this dichotomous empirical gravity equation (40) are estimated through static random effects probit model. It is assumed that there is no state dependence and no serial correlation in the transitory errors. This assumption is clearly unrealistic as cross-border investments are highly dependent on their own past values as pointed out by Drakos et al. (2014). However, a more realistic dynamic random effects probit model is beyond the scope of this thesis.

3.3 Combining the two approaches: a double-hurdle model

The classic empirical gravity model clearly fails to explain the prevalence of zero investments observed as the issue is simply ignored in this approach. As pointed out by Heckmann (1979), treating the zero observations inappropriately may lead to biased and inconsistent estimates. Because the share of zero observations is not only non-negligible, but strikingly high in cross-border investments (see e.g. Drakos et al., 2004), standard regression analysis should not be considered an appropriate approach. Instead, statistical models for limited dependent variable, such as tobit or probit models, should be considered the right approach. On the other hand, the dichotomous approach presented in previous section, though giving due attention to zero holdings, answers only half of the question and loses valuable information by ignoring the levels of cross-border investments. Combining the classic and dichotomous approaches should provide a more thorough and complete description of the determinants of cross-border asset flows. The double-hurdle model first presented by Cragg (1971) and developed further by Heckmann (1976) could provide the combining approach needed to study the determinants of cross-border asset flows in a more realistic manner.

In addition to dealing properly with zero observations, there is another reason that makes the use of the double-hurdle model appealing. While estimating the gravity model by the classical and the dichotomous approach, one might end up with results pointing to different directions. This could mean that there are different processes determining the dichotomous decision of investing or not, and the decision of how much to invest. With the double-hurdle model it is possible to estimate the equation of the dichotomous participation decision and the equation of the continuous level decision simultaneously while allowing them to depend on two different processes (Cragg, 1971, and Heckmann, 1976).

The double-hurdle model is closely related to other limited dependent variables such as the tobit models. As tobit models, double-hurdle models

are used with dependent variables that take a corner solution, such as zero, with a positive, non-negligible probability and are continuously distributed over the set of other solutions. (Garcia, 2013) This is exactly what is observed in the cross-border investments: the distribution of the amounts of foreign assets held are continuously distributed over positive values, but the number of zero holdings is very high and there are no negative investments. Thus the observations pile up at the lower corner of the interval of observed holdings, i.e. at zero.

The tobit model is very commonly used in these kinds of problems, but the functional form of a tobit model restricts the underlying stochastic process to be the same for when the dependent variable equals zero and when it is positive. The double-hurdle model relaxes this assumption and allows these processes to be different. Specifically, the double-hurdle model also allows that the effect of a determinant on the probability that the dependent variable is positive has a *different sign* than the effect of the *same* determinant on the level of the dependent variable, given that it is positive. (Cragg, 1971)

In his seminal paper Cragg (1971) presents several statistical models for limited dependent variables, one of which is the double-hurdle model. All the presented models are based on probit analysis model where the probability that a particular event will occur at period t is

$$P(E_t) = \int_{-\infty}^{X_t'\beta} (2\pi)^{-\frac{1}{2}} \exp\left\{\frac{-z^2}{2}\right\} dz, \quad (41)$$

where X_t is a vector of the independent variables at period t and β is a vector of coefficients. The cumulative unit normal distribution is depicted by

$$C(z) = \int_{-\infty}^z (2\pi)^{-\frac{1}{2}} \exp\left\{\frac{-t^2}{2}\right\} dt. \quad (42)$$

From this starting point Cragg goes on to point out that the tobit model usually used with limited dependent variables will not be appropriate when the decision to acquire and the amount of the acquisition are governed by different effects. To model a situation like this he proposes a model in which

two hurdles must be overcome to achieve positive observations of the dependent variable: First, a decision to consider an acquisition must be made, and second, the amount of the acquisition decided. The density for zero values of the dependent variable y_t is given by

$$f(y_t = 0 \mid X_{1t}, X_{2t}) = C(-X'_{1t}\beta), \quad (43)$$

where X_{1t} and X_{2t} are vectors of independent variables at period t , which are not necessarily distinct, and β is a vector of coefficients. This *participation* decision can be represented by a probit model. The density for positive values of the dependent variable is given by

$$f(y_t \mid X_{1t}, X_{2t}) = (2\pi)^{-\frac{1}{2}}\sigma^{-1}\exp\left\{\frac{-(y_t - X'_{2t}\gamma)^2}{2\sigma^2}\right\}C(X'_{1t}\beta) \quad (44)$$

for $y_t \neq 0$, and where β and γ are vectors of coefficients. This *level* decision can be represented by a standard regression model. It is quite common to restrict y_t to be non-negative by truncating the distribution at zero. The density function related to level decision is then represented by

$$f(y_t \mid X_{1t}, X_{2t}) = (2\pi)^{-\frac{1}{2}}\sigma^{-1}\exp\left\{\frac{-(y_t - X'_{2t}\gamma)^2}{2\sigma^2}\right\}C(X'_{1t}\beta) C\left(\frac{X'_{2t}\gamma}{\sigma}\right) \quad (45)$$

for $y_t > 0$. This representation is appropriate in the case of asset holdings, which are non-negative. The model is estimated by calculating maximum likelihood estimates iteratively, as there is no direct way to solve the first-order conditions from the highly non-linear equations.

The standard double-hurdle assumes bivariate normality of unobserved errors (see e.g. Jones and Yen, 2000), meaning that the errors are independent if and only if there is zero correlation between the errors. In reality the errors may be non-normally distributed resulting to the maximum likelihood estimators being inconsistent. Jones and Yen (2000) solve this problem by generalizing the double-hurdle model to account for dependence between participation and level decisions. They introduce a Box-Cox transformation (Box and Cox, 1964) of the dependent variable, which relaxes the normality assumption on the conditional distribution of the dependent variable and

allows stochastic dependence between the errors. The implementation of Garcia (2013) follows Jones and Yen (2000) and also allows the errors to be correlated. There are several representations in the literature of the estimable decision and level equations (see e.g. Holloway et al., 2005, Blundell et al., 1987, and Heckmann, 1976), but the most suitable for the present purpose is the formulation employed by Garcia (2013) following Jones and Yen (2000). This formulation can be straightforwardly modified to be used with gravity equations. Importantly, Garcia (2013) also introduces a Stata-command for the double-hurdle model, which makes the estimation of the model, acknowledged to be computationally heavy e.g. Cragg (1971), comparatively effortless.

Following the formulation of Garcia (2013) and Jones and Yean (2000), assume that consumers make their consumption decisions in two steps or hurdles. First, a consumer decides whether he wants to make an acquisition at all, that is, to *participate* in the market. This is called the *participation decision*. After this the customer determines the optimal amount of the acquisition according to his preferences. This second decision is called the *level decision*. If the observed consumption is denoted by y_i , the double-hurdle model can be formulated as:

$$y_i = \begin{cases} \beta x_i + u_i & \text{if } \min[\beta x_i + u_i, \gamma z_i + v_i] > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (46)$$

$$\text{where } \begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N(0, \Sigma), \Sigma = \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{21} & 1 \end{pmatrix}.$$

Above $\beta x_i + u_i$ represents the level equation and $\gamma z_i + v_i$ the participation equation. Errors u_i and v_i of the participation and level equation are assumed jointly normal and they may be correlated. (Jones and Yen, 2000, and Garcia, 2013)

When $\Psi(x, y, \rho)$ is the standard bivariate normal cumulative distribution function with correlation ρ , $\phi(x, y, \rho)$ the univariate standard normal density function and $\Phi(x, y, \rho)$ the univariate standard normal distribution function,

the log-likelihood function of the double-hurdle model is

$$\begin{aligned} \log(L) = & \sum_{y_i=0} \left\{ \log \left[1 - \Psi \left(\gamma z_i - c, \frac{\beta x_i - c}{\sigma}, \rho \right) \right] \right\} \\ & + \sum_{y_i>0} \left\{ \log \left[\Phi \left(\frac{\gamma z_i - c + \frac{\rho}{\sigma}(y_i - \beta x_i)}{\sqrt{1 - \rho^2}} \right) \right. \right. \\ & \left. \left. - \log(\sigma) + \log \left[\phi \left(\frac{y_i - \beta x_i}{\sigma} \right) \right] \right] \right\}, \end{aligned} \quad (47)$$

where c is the lower corner, or the point of truncation of the double-hurdle model, which is often and also here assumed to be zero. When $\rho = 0$ and γz_i tends to infinity, the double-hurdle model reduces to a tobit model. (Jones and Yen, 2000, and Garcia, 2013)

The economic interpretation of the estimation results is conducted based on the marginal effects of independent variables on three quantities. First, the probability of the dependent variable being positive:

$$P(y_i > c \mid x_i, z_i) = \Phi \left(\gamma z_i - c, \frac{\beta x_i - c}{\sigma}, \rho \right). \quad (48)$$

Second, the expected level of the dependent variable conditional on the independent variables and the dependent variable being positive:

$$\begin{aligned} E(y_i \mid x_i, z_i, y_i > c) &= \int_c^\infty y_i f(y_i \mid v_i > c - \gamma z_i, u_i > c - \beta x_i) dy, \quad (49) \\ \text{where } f(\cdot) &= \frac{\phi \left(\frac{y_i - \beta x_i}{\sigma} \right) \Phi \left(\frac{\gamma z_i - c + \frac{\rho}{\sigma}(y_i - \beta x_i)}{\sqrt{1 - \rho^2}} \right)}{\sigma \Psi \left(\gamma z_i - c, \frac{\beta x_i - c}{\sigma}, \rho \right)}. \end{aligned}$$

Third, the expected level of the dependent variable conditional only on the independent variables:

$$\begin{aligned} E(y_i \mid x_i, z_i) &= c[1 - P(y_i > c \mid x_i, z_i)] \\ &+ P(y_i > c \mid x_i, z_i)E(y_i \mid x_i, z_i, y_i > c). \end{aligned} \quad (50)$$

The first effect is related to the participation equation and the second effect to the level equation. In the present context c , the lower corner of the model, is assumed to be zero. (Jones and Yen, 2000, and Garcia, 2013)

In the context of cross-border investments, consumers are considered investors and acquisitions in a given market are considered investments into a foreign country. When the double-hurdle model is formulated in the gravity framework:

$$\log T_{ij,t} = \begin{cases} L & \text{if } \min(L, P) > 0, \\ 0 & \text{otherwise,} \end{cases} \quad (51)$$

$$L = \alpha_1 \log M_{i,t} + \alpha_2 \log M_{j,t} + \alpha_3 \log \tau_{ij,t} + u_{ij,t}, \quad (52)$$

$$P = \gamma_1 \log M_{i,t} + \gamma_2 \log M_{j,t} + \gamma_3 \log \tau_{ij,t} + v_{ij,t}, \quad (53)$$

where L is the level equation, P is the participation equation, $T_{ij,t}$ refers to the investments from source country i to destination country j , $M_{i,t}$ and $M_{j,t}$ are measures of the economic masses of country i and j , $\tau_{ij,t}$ denotes transaction costs between the countries. $u_{ij,t}$ and $v_{ij,t}$ are error terms.

As in the classic and dichotomous approaches, the effect of transaction costs can be disaggregated and approximated by distance as a proxy of information frictions, and information variables and transaction technology variables. Time dummies and country dummies are added to empirical equation to capture the effect of macroeconomic impacts and the multilateral resistance term. The double-hurdle model takes the form:

$$\log T_{ij,t} = \begin{cases} L & \text{if } \min(L, P) > 0, \\ 0 & \text{otherwise,} \end{cases} \quad (54)$$

$$\begin{aligned} L = & \alpha_1 \log(M_{i,t}) + \alpha_2 \log(M_{j,t}) + \alpha_3 \log(\text{dist}_{ij}) \\ & + \alpha_4 \log(\text{information variables}) \\ & + \alpha_5 \log(\text{transaction technology variables}) \\ & + \text{time dummies} + \text{country dummies} \\ & + \text{constant} + u_{ij,t}, \end{aligned} \quad (55)$$

$$\begin{aligned}
P = & \gamma_1 \log(M_{i,t}) + \gamma_2 \log(M_{j,t}) + \gamma_3 \log(dist_{ij}) \\
& + \gamma_4 \log(\text{information variables}) \\
& + \gamma_5 \log(\text{transaction technology variables}) \\
& + \text{time dummies} + \text{country dummies} \\
& + \text{constant} + v_{ij,t},
\end{aligned} \tag{56}$$

where L is the level equation, P is the participation equation, $T_{ij,t}$, $M_{i,t}$ and $M_{j,t}$ are as before and $dist_{ij,t}$ refers to the distance between the capitals of the two countries. Also in the double-hurdle model the persistence in observed values of external asset holdings is ignored, and this assumption is acknowledged to be simplifying.

There is no reason why the decision to acquire and the amount of acquisition should be governed by the same effects, that is for $\beta x_i + u_i = \gamma z_i + v_i$ in (46). In many studies these effects have been found to be different and even the coefficients of same determinants to be of different sign, with the classic example being the effect of education on consumption of cigarettes (see e.g. Jones, 1989). In more recent example Garcia (2013) finds that education has a negative effect on the participation equation, but a positive effect on the level equation. That is, education reduces the probability of an individual being a smoker, but given that the individual is a smoker, education correlates positively with the amount of cigarettes he smokes¹¹.

It is not very hard to think of the determinants of the participation decision and the level decision to differ in the context of cross-border investments. It may be that it simply is not possible for residents of source country to invest in a given country. The destination country may not allow foreign investments at all or the government of the source country may have imposed financial sanctions on the destination country that inhibit investments from residents of the source country. These kinds of restrictions might affect only

¹¹For other applications of the double-hurdle model see e.g. Blundell et al., 1987, who study female labor supply, the standard application of the double-hurdle model, Newman et al., 2003, for a model of Irish household expenditure on prepared meals, and Aristei, 2013, for an application considering the remittance behavior of migrants.

the participation equation and result in zero investments. The coefficients of determinants that do effect both equations, such as financial market development or market size in the destination country, may well be different. A bigger market might make the probability of a positive investment larger, but once the decision to invest has been made, the positive effect of market size could be even larger.

The most interesting case of the double-hurdle estimation, the different signs of coefficients of the same determinant, is also thinkable. Distance might affect negatively the probability of making a positive investment; investors may feel that investing into a distant country holds more risk through information frictions than investing on the same continent. But once the positive investment decision is made, investors may consider geographical diversification in the set of countries they did decide to invest in. Then the effect of distance on the level of investment could be positive.

Given the prevalence of zeros in the observed cross-border financial asset holdings and the fact that the two hurdles of cross-border investment decision plausibly could and even probably do depend on different processes, the double-hurdle model should be considered as a correct empirical approach to estimating the gravity equation of cross-border investments. I illustrate my argument in the next section with the help of Finnish external assets.

4 Empirical application: the external wealth of Finland

The external wealth of Finland is well characterized by the two phenomena of cross-border investments described in the introduction: home bias and prevalence of zero observations in cross-border investments. According to the official statistics, Finnish investors invest relatively little, on average 46% in 1998-2013 of portfolio assets in equities and debt securities (see figure 1, data from Bank of Finland and Statistics Finland), of their total wealth abroad. However, the share of the Finnish economy from the world economy is very small: GDP of Finland accounted for about 0.4% of the world GDP in 2013 according to IMF. It seems that Finns suffer from a serious case of home bias and care very little for geographical diversification and risk sharing.

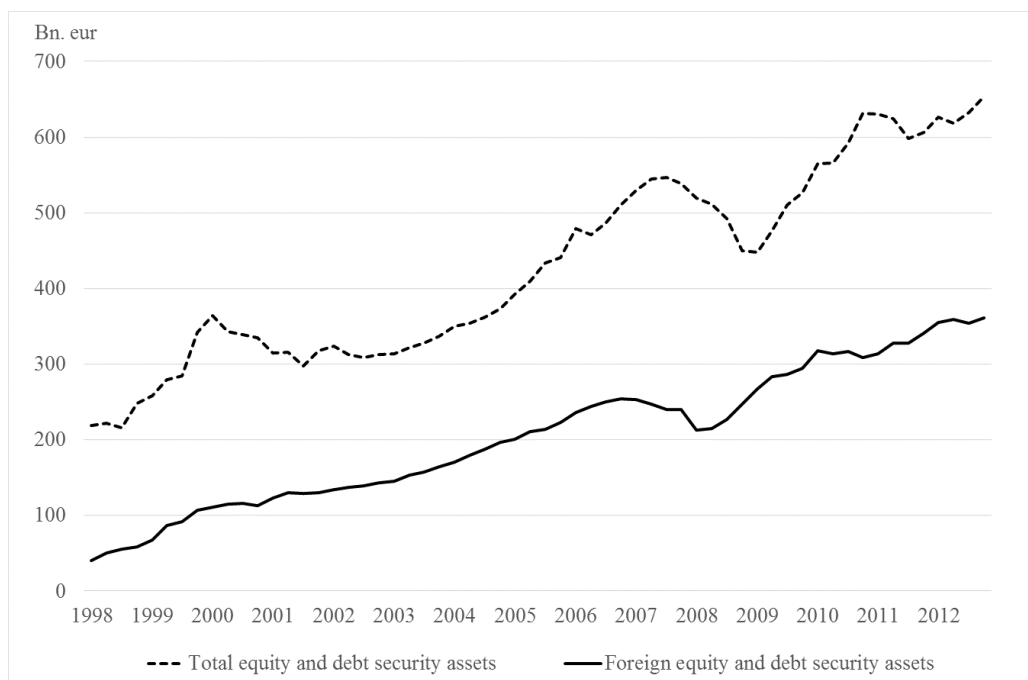


Figure 1: Total and foreign equity and debt security assets held by Finnish residents

The dominance of zero holdings is also evident in the external wealth of

Finland. Of all annual observations in 2004-2013 on cross-border holdings of equity 64.27% and of debt securities 66.18% were zero (Bank of Finland). That is, Finnish investors invest in only a fairly small number of certain other countries and the distribution of investments across destination countries is clustered on zero. Even though this concentration of investments has been on decline over the years (see table 1, data from Bank of Finland), it is still a very prominent feature of the cross-border investments of Finnish residents.

	equities	debt securities
2004	73.44%	74.69%
2005	70.95%	74.27%
2006	65.15%	73.44%
2007	61.41%	65.98%
2008	63.90%	66.39%
2009	63.90%	65.98%
2010	63.90%	63.49%
2011	63.49%	62.66%
2012	58.51%	59.34%
2013	58.09%	55.60%

Table 1: Share of zero holdings in observed cross-border investments of Finnish residents

The data on Finnish external wealth shows that the geographical distribution of external assets held by Finnish investors is skewed: Finnish investors invest much more e.g. in Sweden compared to USA, even though USA is a much larger economy with a business cycle much less correlated with that of Finland. In 2013 the geographical distribution of Finnish portfolio assets was very limited, with domestic assets accounting for almost 40% of the assets and assets of only five countries accounting for the next 35% (see figure 2, data from Bank of Finland¹²). There in fact appears to be a strong negative

¹²The large share of Luxembourg is explained by Luxembourg's sizeable mutual fund

correlation between bilateral distance and cross-border assets held by Finnish residents both in equities and debt securities in 2004-2013 (see figure 3, data from Bank of Finland and CEPII).

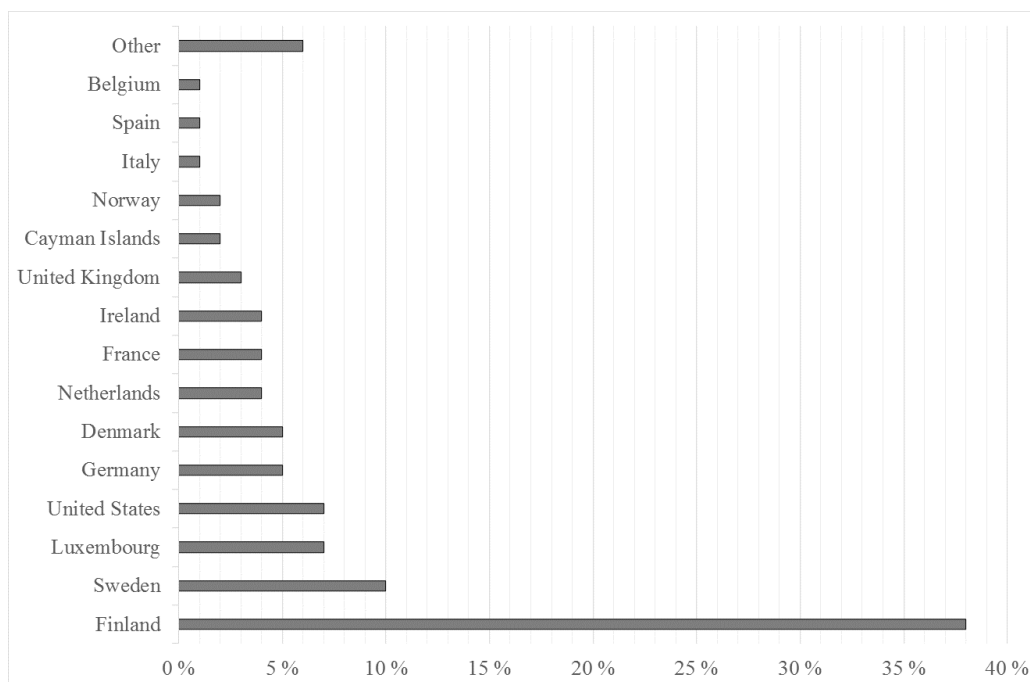


Figure 2: The geographical distribution of assets held by Finnish residents in 2013

As Finnish external assets display neatly the two prominent phenomena of cross-border asset holdings, the case of Finnish external wealth interesting to study in the gravity framework. There is however a caveat in the data on external wealth. External assets are notoriously difficult to capture in statistics (see e.g. Zorzi, 2009, and Zucman, 2013) due to the tendency to underreport assets in order to minimize or evade taxes, a phenomenon familiar in all countries. In addition to this, there is a problem of limited coverage in the surveys which are used to compile data on external assets. In Finland, industry, which results in the statistics being biased as the investments' final destination is some other country. For a thorough consideration of the caveats of the CPIS data, see Felettigh and Monti, 2008.

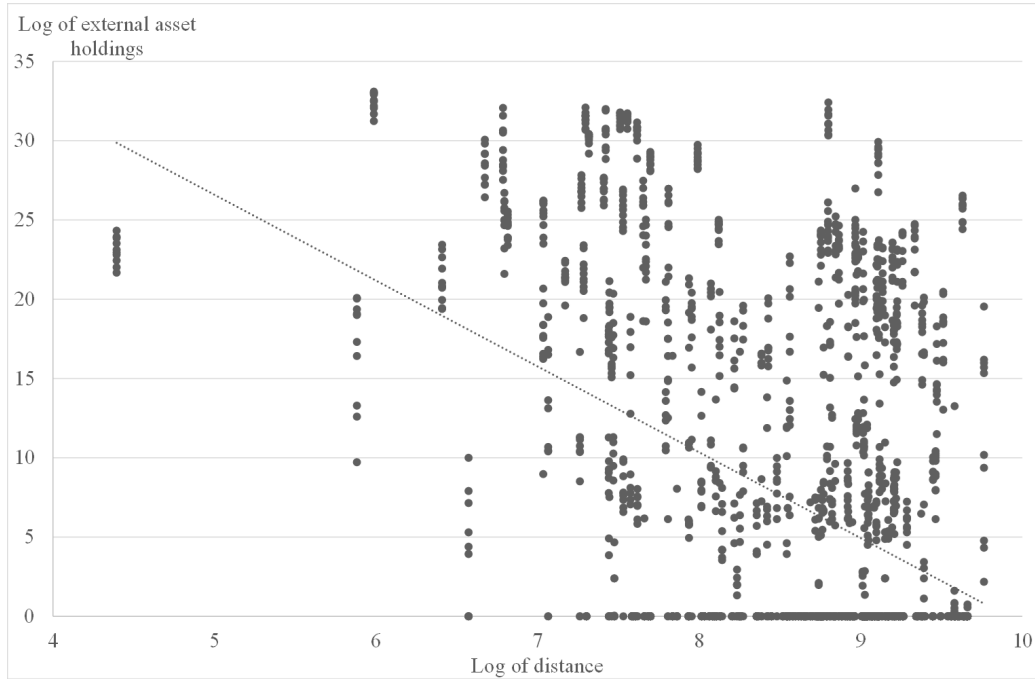


Figure 3: Correlation of distance and external portfolio assets held by Finnish residents

all institutional investor sectors are relatively well covered in these surveys, but the cross-border investments made by certain sectors, e.g. households, without using a *domestic* intermediate are not captured by the statistics. Thus the level of external assets held by Finnish residents according to official statistics is underestimated and acknowledged to be so¹³. This may result in a more sizeable home bias, more zero holdings and a more concentrated geographical distribution than reality would allow. However, the limitations of the official statistics are highly unlikely to account for the observed features of the external wealth of Finland completely. Thus studying the determinants of cross-border investments of Finnish residents with the gravity framework will serve to illustrate the differences between the different empirical approaches to estimating gravity equations.

¹³The interested reader may refer to Salo, 2014, for a more extensive discussion on the problems related to data quality.

4.1 Data on Finnish external wealth and its determinants

The data is a panel of 241 destination countries and years 2004-2013, and there are $10 * 241 = 2410$ observations of the dependent variables. All data are in logs or dichotomous. In the destination countries I include all 241 of the possible destination countries listed in the Coordinated Portfolio Investments Survey or CPIS-data of IMF (excluding Finland, which is the source country). The data on external assets of Finnish residents is obtained from International Investment Position (IIP) statistics of Finland (surveyed and compiled by Bank of Finland). In the classic approach to estimating the gravity model the holdings of financial assets are in logs (eq_{ij} denotes equity holdings and de_{ij} debt security holdings), and in the dichotomous approach they take value 0 if the stock of equity is zero and value 1 if it is positive (dq_{ij} denotes dichotomous equity holdings and dde_{ij} dichotomous debt security holdings). The summary statistics on the dependent variables are given in table 2.

Variable	Mean	Std. dev.	Min	max	Obs.
eq_{ij}	3.69	5.37	0	17.12	2410
de_{ij}	3.51	5.24	0	17.12	2410
dq_{ij}	0.36	0.48	0	1	2410
dde_{ij}	0.34	0.47	0	1	2410

Table 2: Dependent variables

The economic masses are measured by gross domestic products (in current euros), obtained from the World Bank Economic Indicators. In addition to GDP, GDP per capita and market capitalization are often used as measures of economic mass. However, to my understanding GDP provides the best measure for economic mass of these three options, because it describes the economic significance of a country as a whole, instead of simply describing

the average income of its residents as the GDP per capita or the size of the stock markets as the market capitalization. Thus the variables describing economic mass are gdp_i and gdp_j , where i stands for source country and j for destination country. I interpret transaction costs broadly to include information costs and transaction costs that can be interpreted also as measures of attractiveness to investment. Distance ($dist_{ij}$), is considered to proxy information frictions. Bilateral distance is measured as the theoretical air distance (great circle distance) between the capital of the source country and the capital (or major city if there is now official capital) of the destination country (obtained from CEPII).

To measure information advances Portes and Rey (2005) use telephone traffic, number of bank subsidiaries, and degree of overlap in trading hours and index on insider trading in the destination country. They also use indicator variables for contiguity and same official language. Finland, being a relatively small country, has only a few bank subsidiaries situated in other countries and has a common official language with only Sweden. Common language and contiguity did not appear to work well in the preliminary estimations, due to collinearity perhaps, so they are omitted. I employ an indicator variable describing whether a Finnish bank has a subsidiary in the destination country ($banksubs$)¹⁴. The other three variables employed by Portes and Rey (2005), telephone traffic, common trading hours and index of insider trading, I found to be unavailable. I believe however that at least having to do without telephone traffic data should not be of grave consequences. Telephone traffic, although found to have some significance in their sample dated in 1989-1996, is not likely to have that much significance during my sample period of 2004-2013, with the internet having replaced telephone calls as the primary source of information. Search engine searches or the number of emails sent from domestic servers to foreign ones might provide a more

¹⁴This could also be an absolute number, but as the maximum number of foreign subsidiaries of Finnish banks is only two (in Sweden, Estonia, Lithuania and Latvia), this variable can easily be treated as dichotomous.

sensible proxy for international information flows. The degree of overlap in trading hours could be significant still, but even Portes and Rey (2005) employ it only in their robustness checks and find that even though the variable is significant, it does not alter their main results. Additionally the authors found insider trading to be somewhat unstable, and to have only minor effects on the overall results. Thus omitting these three information variables should not compromise the results too much. In addition to distance and common language, Drakos et al. (2014) employ common legal origin as a measure of information flows, as common legal origin should lower the costs of using information relevant for investment decisions. I use a dummy variable taking value 1 if the legal origin of the destination country is the same, "Scandinavian civil law", as it is in Finland (*lawscan*). In addition to the variables mentioned above, I also include a dummy variable for European Union membership (*eu*), to describe information effects.

In the measures of transaction costs or attractiveness of investment Portes and Rey (2005) and Drakos et al. (2014) include different measures of financial market development, such as the domestic credit per GDP, the stock turn-over ratio or a specific index of financial market sophistication. I use domestic credit to GDP $domcred_j$ available from World Bank to proxy financial sophistication, but not stock turn-over ratio, because this data is available for relatively few countries. In addition, I include an indicator variable for euro area membership (*euro*). Openness of the destination country, measured as exports and imports per GDP (exp_j and imp_j), is thought to proxy the ease at which trade in financial assets in the destination country is carried out.

The role of bilateral goods trade is found to be somewhat ambiguous in the literature. Portes and Rey (2005) study the effect of trade, but found their main result to hold. Aviat and Couerdacier (2007) on the other hand argued that there is complementarity between bilateral trade in goods and bilateral asset holdings. To my understanding bilateral trade can also be thought to effect information flows. On the import side this could be be-

cause consuming products of a certain country can be thought to be closely connected to the degree of familiarity of the culture of that specific country. On the export side, it could be argued that exporting to a certain country increases the amount of aggregate knowledge of the destination country in the exporting country. Exporting firms must have sophisticated knowledge of the destination economy in order to be successful exporters, and this information might be thought to leak into the society as a whole. In order to study the effects of trade, I include the bilateral trade flows calculated as the sum of exports and imports between the source country and the destination country (*trade*). These data are obtained from the Balance of Payments statistics of Finland (Bank of Finland). The summary statistics on the independent variables are given in table 3.

	Variable	Mean	Std. dev.	Min	Max	Obs.
Economic masses	gdp_i	25.47	1.28	20.78	25.98	2410
	gdp_j	23.19	2.68	12.07	30.17	1868
Information effects	$dist_{ij}$	8.58	0.81	4.39	9.76	2410
	$banksubs$	0.04	0.19	0	1	2410
	$lawscan$	0.02	0.13	0	1	2410
	eu	0.11	0.31	0	1	2410
Transaction effects	$euro$	0.06	0.23	0	1	2410
	$domcred_j$	3.63	0.93	0.60	5.77	1531
	imp_j	3.82	0.48	0.12	5.44	1646
	exp_j	3.64	0.59	0	5.44	1647
Bilateral trade	$trade$	8.51	4.60	0	16.93	2410

Table 3: Independent variables

To control for multilateral resistance term I include regional dummies for North America, South America, Europe, Asia, Africa and Oceania. Time specific effects are included in all estimations to capture the effect of macroeconomic shocks.

4.2 Methods

The simplified theoretical gravity equation (36) is taken as a starting point and the full model to be estimated is

$$\begin{aligned}
 y_{ij,t} = & \beta_1 gdp_{i,t} + \beta_2 gdp_{j,t} + \beta_3 dist_{ij,t} + \beta_4 banksubs + \beta_5 lawscan \\
 & + \beta_6 eu + \beta_7 domcred_j + \beta_8 euro + \beta_9 imp_j + \beta_{10} exp_j + \beta_{11} trade \\
 & + \text{time dummies} + \text{regional dummies} + \text{constant} + u_{ij,t}, \quad (57) \\
 & i, j = 1, \dots, 241 \text{ and } t = 1, \dots, 10.
 \end{aligned}$$

The dependent variable $y_{ij,t}$ is the level of equity or debt security holdings, eq_{ij} and de_{ij} , in the classic approach and in the level equation of the double-hurdle model. In the dichotomous approach and the participation equation of the double-hurdle model the dependent variable $y_{ij,t}$ is dichotomous equity or debt security holdings, deq_{ij} and dde_{ij} , which take value 0 if a zero holding is observed and value 1 otherwise. Recall from section 3.1 that theory predicts coefficients β_1 and β_2 to be close to 1, β_3 to be negative and the rest of the coefficients to be positive.

There is only one source country, as the consideration is limited to investments of Finnish residents. This leads to there being a smaller amount of observations compared to the estimations of Portes and Rey (2005) and Drakos et al (2014) who have 14 and 58 source countries respectively. Additionally, restricting the number of source countries to only one makes the variation in the independent variables over time and over source-destination-country pairs limited. For example, the economic mass of the source country changes only over time and even then fairly little, the changes in GDP being in general slow and not dramatic. In addition, many of the controls are closely connected and there is a lot of covariance between the independent variables.

The limited variation and collinearity pose some problems to the estimation methods employed. Portes and Rey (2005) use fixed effects estimation, but due to the above mentioned problems, straightforward fixed effects estimation results in most of the independent variables being omitted due to

collinearity. Using random effects estimation would not result in omitting important variables, but according to the Hausman-test the random effect estimates would be inconsistent, and thus random effects estimation is an inappropriate method in this case. To account for the omitted variables the estimation is done as follows:

1. The fixed effects estimation is performed for all included variables.
2. The country specific fixed effects are calculated using *predict* post-estimation command.
3. The calculated country specific fixed effects are regressed on the omitted variables.
4. The marginal effects of the independent variables on the level of the dependent variable are calculated.

To explain the prevalence of zero observations, a dichotomous model is estimated. This is done following the static model of Drakos et al. (2014). The estimations are performed as follows:

1. The random effects probit model for panel data is estimated.
2. The marginal effects of the explanatory factors on the probability of there being a positive investment are calculated.

In addition to presenting the results of these two empirical models in the context of cross-border investments of Finnish residents, I propose a method that has previously not been applied to estimating the determinants of cross-border investments: The double-hurdle model. The double-hurdle estimations are done following Garcia (2013):

1. The double-hurdle model is estimated.
2. The marginal effects of the explanatory factors on the probability of there being a positive investment are calculated.

3. The marginal effects of the independent variables on the level of investment *given that the investment is positive* are calculated.

One should note that the first two models are for panel data, but the double-hurdle model is not and instead considers the observations simply pooled. However, as pointed out in section 4.3, there are significant individual effects in the data and thus considering the observations pooled may not be appropriate. This should be kept in mind while interpreting the results.

The first specification includes only GDP's approximating the economic masses of the source and destination countries and distance approximating information frictions. In the second specification different variables proxy the effect of information and transaction costs. In the third specification bilateral trade and multilateral resistance in the form of regional dummies are controlled. All equations include a constant term and time dummies, estimates of which are not reported. All variables are in logs or dichotomous.

4.3 The classic approach: the determinants of the level of cross-border investments

The results of the fixed effects and cross-section estimations are given in tables 4 and 5 for equity and debt security holdings, with standard errors in parentheses. In table 4 I report the marginal effects of independent variables on the level of dependent variables and in table 5 other statistics. Specifications (1), (2) and (3) have equity holdings as the dependent variable, while (4), (5) and (6) have debt security holdings. The results are for the most part in accordance with previous studies and confirm e.g. the negative effect of distance on both equity and debt security holdings. There is fairly little difference between the results for determinants of equity holdings and debt security holdings. Most of the statistically significant variables remain so throughout the specifications and enter all specifications with the same signs and order of magnitudes.

The overall fit of the model is not terrific for either of the dependent variables (see table 5), but it is somewhat increased with the inclusion of additional information and transaction variables in specifications (2) and (5). Including trade and the regional dummies in specifications (3) and (6) has virtually no effect. The overall coefficients of determination (R^2) of the fixed effects regression fall between 0.41 and 0.58 with both equities and debt securities as dependent variables. For the cross-section regressions the usual and adjusted coefficients of determination fall between 0.09 for debt security holdings and 0.25 for equity holdings as dependent variables. The degree of variation in dependent variable due to residual $u_{ij,t}$, ρ , is quite stable throughout the specifications, varying from 0.76 to 0.86. ρ can be interpreted as the share of variation related to differences in destination countries and thus it being quite high is not surprising.

All specifications were also estimated by the random effects model, but in all cases the Hausman test indicated that the random effects estimators were inconsistent and thus the fixed effects estimators were the appropriate ones. The F-tests for individual effects indicate that in all specifications

there are significant individual effects and thus pooled ordinary least squares estimation would be inappropriate.

The GDPs, in the role of the economic masses, do not behave fully as expected: recall from section 3.1 that theory predicted the coefficients of economic masses to have a coefficient close to one. The effect of GDP of destination economy is highly significant and has a relatively large positive effect, from 0.72 to 2.53, on the dependent variables. This is as expected; a larger economy attracts more foreign investments. The effect is however clearly larger in debt securities, and for equities the size of the destination country seems to play a smaller role. Surprisingly, the GDP of the investing country has a fairly large negative effect, from -2.42 to -0.23, on both of the dependent variables, however being statistically significantly different from zero only with debt securities. I think this could be explained by the strong home bias of Finnish investors: when domestic GDP is growing, investing home seems even more tempting, thus reducing the level of cross-border investments. However, the GDP of the source country varies here very little across time and country pairs and this may affect the results.

The effect of distance is exactly as expected: negative and strongly statistically significant throughout the specifications. The effect is almost halved by the additional variables in specifications (2), (3), (5) and (6), but remains nevertheless quite large. The negative effect is larger in equities, from -2.1 to -1.1, than in debt securities, from -1.24 to -0.72. This is not surprising: It is easy to think that investing in equities is more information intensive, whereas debt securities are less risky, more standardized and investing in them might require less specific information.

Of the variables proxying information effects the indicator variables for bank subsidiaries and Scandinavian law system both have large, positive and statistically significant coefficients in all specifications. This might be related to the fact that Sweden receives the largest shares of Finnish cross-border investments. The indicator variable for European Union membership somewhat surprisingly has a negative effect, which is statistically significant

for debt securities, but not for equities.

Of the variables proxying transaction effects only the indicator variable for eurozone has a significant effect throughout the specifications, being larger for equities than for debt securities. In other variables there are differences between the dependent variables, e.g. the share of domestic credit of GDP surprisingly has a negative, statistically significant effect on debt security holdings, but no statistically significant effect on equities. The effect of openness of the destination economy remains also small and less significant.

Controlling for bilateral trade flows and multilateral resistance in the form of regional dummies has no effect: none of these variables has a statistically significant coefficient.

From these results it seems that the level of cross-border investments of Finnish residents is mainly determined by the size of the destination economy and the bilateral distance with GDP of the destination economy having a positive effect and distance a negative effect, as expected. Also the destination country having Scandinavian law system, hosting subsidiaries of Finnish banks and belonging to the euro zone have positive, large and statistically significant effect. The standard results of distance being negatively correlated with cross-border investments is confirmed in all specifications for both equity and debt security holdings.

	eq_{ij} (1)	eq_{ij} (2)	eq_{ij} (3)	de_{ij} (4)	de_{ij} (5)	de_{ij} (6)
gdp_i	-1.08**** (0.29)	-0.27 (0.58)	-0.23 (0.58)	-1.27**** (0.33)	-2.42**** (0.67)	-2.33**** (0.67)
gdp_j	0.72*** (0.24)	1.28**** (0.34)	1.24**** (0.35)	1.41**** (0.28)	2.53**** (0.40)	2.43**** (0.40)
$dist_{ij}$	-2.10**** (0.11)	-1.11**** (0.11)	-1.10**** (0.11)	-1.24**** (0.09)	-0.79**** (0.14)	-0.72**** (0.14)
$banksubs$	-	4.13**** (0.40)	4.13**** (0.40)	-	2.80**** (0.52)	2.75**** (0.53)
$lawscan$	-	3.44**** (0.58)	3.40**** (0.58)	-	6.49**** (0.75)	6.43**** (0.76)
eu	-	-1.44 (1.17)	-1.44 (1.18)	-	-3.00** (1.36)	-2.99** (1.36)
$domcred_j$	-	0.20 (0.26)	0.20 (0.26)	-	-0.93*** (0.30)	-0.97*** (0.30)
$euro$	-	3.29**** (0.62)	3.62**** (0.62)	-	1.72** (0.72)	1.67** (0.72)
imp_j	-	0.28 (0.44)	0.27 (0.44)	-	0.95* (0.50)	0.90* (0.51)
exp_j	-	0.66* (0.39)	0.67* (0.39)	-	-0.33 (0.45)	-0.34 (0.45)
$trade$	-	-	0.03 (0.05)	-	-	0.09 (0.06)
regional	No	No	Yes	No	No	Yes

Significance at the 10%, 5%, 1% and 0.1% levels is denoted by *, **, *** and **** respectively.

Standard errors in parentheses.

Table 4: Marginal effects of the independent variables on the level of cross-border equity and debt security holdings in the classic approach

eq_{ij}	(1)	(2)	(3)
N	1868	1455	1455
Fixed effects regression			
R^2 , overall	0.46	0.58	0.58
F-test statistic for individual effects	12.78	10.39	7.66
ρ , fraction of variance due to $u_{ij,t}$	0.86	0.81	0.81
F-test statistic for $u_i = 0$	47.73	21.35	21.29
Cross-section regression			
F(K, N-K-1)	379.76	163.08	160.62
R^2	0.17	0.25	0.25
R^2 , adjusted	0.17	0.25	0.25
de_{ij}	(4)	(5)	(6)
N	1868	1455	1455
Fixed effects regression			
R^2 , overall	0.50	0.41	0.41
F-test statistic for individual effects	24.46	15.74	11.43
ρ , fraction of variance due to $u_{ij,t}$	0.76	0.83	0.83
F-test statistic for $u_i = 0$	30.64	13.66	13.57
Cross-section regression			
F(K, N-K-1)	179.89	81.58	74.06
R^2	0.09	0.14	0.13
R^2 , adjusted	0.09	0.14	0.13

Table 5: Estimation results for the level of cross-border equity and debt security holdings

4.4 The dichotomous approach: the determinants of the probability of a positive cross-border investment

The results of the random effects probit estimation are given in tables 6 and 7 for the dichotomous cross-border equity and debt security holdings, with standard errors in parentheses. Specifications (1), (2) and (3) have equity holdings as the dependent variable, while (4), (5) and (6) have debt security holdings. The resulting marginal effects in the probit model are not *per se* comparable to the marginal effects calculated from the fixed effects model in previous subsection 4.3. It is however possible to point out some qualitative similarities and differences between the models.

As in the classic approach, here too there is fairly little difference between the results for determinants of the probability of positive equity or debt security holdings. The degree of variation in dependent variable due to residual $u_{ij,t}$, ρ , is quite high, ranging from 0.82 to 0.92. It thus appears that the fit of this model is not terrific either. This statistic is also quite stable throughout the specifications, but it is clearly higher in the first specifications (1) and (4), than after the inclusion of information and transaction effects in specifications (2) and (5). Including bilateral trade and the regional dummies has no effect with equities in specification (3), but with debt instruments in specification (6) ρ is lower. The Wald test statistics imply that the variables are strongly jointly significant across all six specifications.

Due to the non-linear nature of the probit model and the limited variance and collinearity of the independent variables, the estimation is not possible without omitting some of the indicator variables. With equity holdings two indicator variables, for bank subsidiaries and eurozone, are omitted, and with debt security holdings the indicator variable for Scandinavian law system is omitted.

The effect of GPD of the destination country is as expected, with positive and statistically significant effects in all specifications, in accordance with

the classic approach. The effects appear small in magnitude, ranging from 0.08 to 0.13, but they are statistically significantly different from zero. The effect of GDP of the investing country is ambiguous contrast to the classic approach: the effect is negative in specifications (1) and (4), but positive after the additional variables are included, from -0.13 to 0.36. The effect of domestic GDP remains statistically significant throughout specifications.

The effect of distance is not at all as clear as it was in the classic approach. The effect is negative and statistically significant in only the first specifications, (1) and (4): -0.12 with dichotomous equities and -0.05 with debt securities as the dependent variable. As in the classic approach, the effect of distance is less strong for the probability of debt holdings being positive than for the probability of equity holdings being positive. After the additional variables are included, the effect distance ceases to be significantly different from zero (though the effect even turns positive for debt securities). It appears that distance plays a much smaller role in the dichotomous investment decision than in the classic level decision¹⁵.

However, the indicator variables that are supposed to proxy information advances, do not have unambiguously statistically significant effects either, although the effects are positive as expected. The lost significance of distance thus seems to be due to something else than the added information variables. This raises questions about the appropriateness of these variables in proxying information frictions after all, but as mentioned before, with only one investing country, there is very little variance in these variables.

The variables proxying transaction effects are also quite unstable. The share of domestic credit of GDP in the destination economy has a highly significant positive effect, 0.09, on the probability of positive equity holdings, but not on the probability of positive debt security holdings. Openness of the destination economy proxied by share of imports and exports from GDP have mixed effects.

¹⁵This is however also in contrast to Drakos et al., 2014, who confirm the negative effect of distance.

Controlling for bilateral trade flows and the multilateral resistance term have quite small effects on the results, but trade flows does have positive and statistically significant effect, albeit small at 0.02, on the probability of positive debt security holdings. This is contrast to the results of the classic approach, where trade had no effect at all.

From these results it appears that the dichotomous decision of making a cross-border investment of Finnish residents is mainly determined by the size of the destination economy and of the source economy. All other variables, distance included, seem to play only a limited role.

The results of this dichotomous model are interestingly different from the previous classic approach, with only the effect of GDP of the destination economy having similarly positive and statistically significant effect in both models. Effect of domestic GDP is ambiguous and the role of distance is much reduced and more ambiguous in the probit model. The effects of other controls are also mixed. This can be interpreted as pointing out to there being different processes determining the dichotomous decision of investing in a given country and the level of that investment, given it is positive. The differences in the results serve as a motivation for the use of a double-hurdle model.

	<i>deq_{ij}</i> (1)	<i>deq_{ij}</i> (2)	<i>deq_{ij}</i> (3)	<i>dde_{ij}</i> (4)	<i>dde_{ij}</i> (5)	<i>dde_{ij}</i> (6)
<i>gdp_i</i>	-0.13**** (0.02)	0.27** (0.12)	0.26** (0.11)	-0.09**** (0.02)	0.36** (0.14)	0.36** (0.14)
<i>gdp_j</i>	0.11**** (0.00)	0.08**** (0.01)	0.08**** (0.01)	0.13**** (0.01)	0.12**** (0.01)	0.10**** (0.01)
<i>dist_{ij}</i>	-0.12**** (0.02)	-0.04 (0.03)	-0.03 (0.03)	-0.05** (0.02)	0.01 (0.03)	0.02 (0.03)
<i>banksubs</i>	-	omitted	omitted	-	0.38** (0.16)	0.34** (0.16)
<i>lawscan</i>	-	0.11 (0.17)	0.11 (0.17)	-	omitted	omitted
<i>eu</i>	-	0.14* (0.08)	0.14* (0.07)	-	0.01 (0.06)	0.02 (0.06)
<i>domcred_j</i>	-	0.09**** (0.02)	0.09**** (0.02)	-	0.03 (0.02)	0.03 (0.02)
<i>euro</i>	-	omitted	omitted	-	0.06 (0.06)	0.06 (0.06)
<i>imp_j</i>	-	-0.01 (0.05)	-0.09 (0.05)	-	0.10** (0.05)	0.09* (0.05)
<i>exp_j</i>	-	0.07* (0.04)	0.07* (0.04)	-	-0.04 (0.04)	-0.05 (0.01)
<i>trade</i>	-	-	0.00 (0.01)	-	-	0.02** (0.01)
regional	No	No	Yes	No	No	Yes

Significance at the 10%, 5%, 1% and 0.1% levels is denoted by *, **, *** and **** respectively. Standard errors in parentheses.

Table 6: Marginal effects of the explanatory factors on the probability of positive cross-border equity and debt security holdings in the dichotomous approach

deq_{ij}	(1)	(2)	(3)
N	1868	1454	1454
Wald test statistic	164.37	116.37	116.66
Degrees of freedom	12	15	21
ρ , the fraction of variance due to $u_{ij,t}$	0.92 (0.02)	0.85 (0.03)	0.85 (0.03)
dde_{ij}	(4)	(5)	(6)
N	1868	1454	1454
Wald test statistic	173.32	132.83	137.11
Degrees of freedom	12	16	22
ρ , the fraction of variance due to $u_{ij,t}$	0.89 (0.02)	0.82 (0.04)	0.80 (0.04)

Standard errors in parentheses.

Table 7: Estimation results for dichotomous cross-border equity and debt security holdings

4.5 The double-hurdle approach: the determinants of the participation and level decisions

The differences between the results of the two previous approaches suggest that there could be different processes determining the dichotomous decision of investing or not, and the decision of how much to invest. That is, the determinants of the participation equation and the level equation of cross-border equity and debt security holdings may be different. Also the same determinants may have an effect of different magnitude or even of different sign on the two equations. With the case of cross-border holdings of Finnish residents, this appears to be shown already in the qualitatively different results of the classic approach and the dichotomous approach. Thus there is a clear motivation for estimation the determinants of cross-border investments by the double-hurdle model.

The results of the double-hurdle estimation for equity holdings are given in table 8 and for debt security holdings in table 9, with additional statistics for both dependent variables given in table 10.¹⁶ The participation equation, denoted by P, determines the effect of the independent variables to the probability of a positive investment, corresponding to the dichotomous approach. The level equation, denoted by L, determines the effect of the independent variables on the level of the investment, corresponding to the classical approach, *given that the investment is positive*. The specifications are similar to those of the classical and the dichotomous approach regards to the independent variables included. All variables are included in the participation and the level equation, though this need not be, as different variables may have significant effects on the two decisions.

With equity holdings as the dependent variable the main differences between the participation and the level equation are in the magnitudes of the effects, but there are some differences in the levels of statistical significance

¹⁶The first specification with equity holdings as the dependent variable is estimated only with data from years 2004-2007 and 2013. Including the years of the financial crisis resulted in uncalculatable numerical derivatives.

also (see table 8). Most of the statistically significant variables however remain so in both equations of all specifications. The order of magnitudes of the effects are comparable between the participation equation and the dichotomous probit model and on the other hand the level equation and the classic fixed effects approach.

The interesting case of the effect of the same variable being of a different sign on the participation and level equations is displayed right in the first specification with the effect of GDP of the investor country: The effect is positive, 0.38, on the participation equation and negative, -2.84, on the level equation and statistically significant in both. In the following specifications the effect remains positive on participation equation and negative on level equation, but remains statistically significant only in the level equations. This can be interpreted as the GDP of the investor country increasing the probability of a cross-border investment, but decreasing the level a positive cross-border investment. Perhaps this could be explained by home bias again.

The effect of GDP of the destination country is very similar to the results of the classic approach and the dichotomous approach. The effect is positive and highly significant throughout all specifications, ranging from 0.06 to 0.1 on participation equation (the probability of there being a positive investment) and from 0.95 to 1.11 on level equation.

Distance has a negative effect in all equations, the effect on the participation equations ranging from -0.09 to -0.02 and from -0.83 to -0.25 in the participation equations. The effects are statistically significantly different from zero in all but one specification. Of the information effect variables only the indicator variable for Scandinavian law system is statistically significant with positive effect. The effects of EU membership and bank subsidiaries are not significantly different from zero.

Transaction variables are thought to effect the probability of a positive investment and the amount invested through financial market sophistication in the destination economy proxied by domestic credit per GDP and openness of the destination economy proxied by imports and exports per GDP. All of

these have statistically significant effects. Domestic credit has a positive effect as expected, but the effect of imports and exports is a bit puzzling. Imports seem to have a positive effect, while exports have a negative effect. That is, importing countries attract foreign investments while exporting countries do not. If financial asset flows were a mirror image of trade flows, this should be the other way around. Even more puzzling, the share of imports of GDP has the biggest effect on the level equation, even larger than GDP of the destination economy. These results may however be due to some peculiarity in the data.

Including trade and the regional dummies affects the estimates very little. Bilateral trade volume has a statistically significant positive effect on the probability of a positive investment. If interpreted boldly, this might support the idea of trade carrying information effects and thus making cross-border investments more probable.

With the debt instruments as the dependent variable, the differences between the participation and level equations are mainly as above with equity holdings (see table 9). GDP of the investor country has a negative and statistically significant effect in both participation (-0.85) and level equation (-0.13) in the specification (4). When additional variables are added in the two following, the effect is no longer statistically significantly different from zero. The GDP of the destination country enters all equations with positive and highly statistically significant coefficients, ranging from 0.08 to 0.12 on the participation equation and from 0.69 to 0.8 on the level equation. These results are very similar than the ones with equity holdings as dependent variable.

Distance has an as-expected negative effect in both equations in the specification (4), -0.06 on participation equation and -1.03 on level equation, but in the specifications (5) and (6) distance has a positive and statistically significant effect on participation equations, but a negative, though not significant, effect on level equations. Of the variables proxying information advances only shared law system had a statistically significant effect on equity holdings, but

	(1)*		(2)		(3)	
	P	L	P	L	P	L
<i>eq_{ij}</i>						
<i>gdp_i</i>	0.38*** (0.14)	-2.84** (1.44)	0.14 (0.11)	-1.77**** (0.50)	0.13 (0.11)	-1.67*** (0.51)
<i>gdp_j</i>	0.10**** (0.00)	0.95**** (0.06)	0.09**** (0.01)	1.11**** (0.06)	0.06**** (0.01)	1.01*** (0.12)
<i>dist_{ij}</i>	-0.09**** (0.01)	-0.83**** (0.12)	-0.02* (0.01)	-0.32**** (0.11)	-0.02 (0.01)	-0.25** (0.13)
<i>banksubs</i>	-	-	0.78 (16.86)	0.11 (17.97)	0.78 (40.14)	-0.16 (50.93)
<i>lawscan</i>	-	-	0.21** (0.10)	0.91** (0.44)	0.19* (0.10)	0.82* (0.45)
<i>eu</i>	-	-	0.05 (0.04)	0.42 (0.31)	0.03 (0.04)	0.35 (0.31)
<i>domcred_j</i>	-	-	0.10**** (0.01)	0.73**** (0.13)	0.10**** (0.01)	0.73**** (0.14)
<i>euro</i>	-	-	0.74 (16.79)	-0.16 (17.90)	0.75 (32.07)	-0.30 (40.69)
<i>imp_j</i>	-	-	0.11*** (0.03)	2.22**** (0.35)	0.09*** (0.03)	2.16**** (0.35)
<i>exp_j</i>	-	-	-0.04** (0.02)	-0.95**** (0.29)	-0.04* (0.02)	-0.95**** (0.29)
<i>trade</i>	-	-	-	-	0.03*** (0.01)	0.10 (0.10)
regional	No	No	No	No	Yes	Yes

*The first specification is estimated with data only on years 2004-2007 and 2013.

Significance at the 10%, 5%, 1% and 0.1% levels is denoted by *, **, *** and **** respectively.

Standard errors in parentheses.

Table 8: Marginal effects of the independent variables on cross-border equity holdings in the double-hurdle model

with debt securities the result is the opposite. Now Scandinavian law system has no effect, but instead there being a subsidiary of a Finnish bank in the destination country has a statistically significant effect and positive (from 0.45 to 0.49) effect on the participation effect. EU-membership has a large, statistically significant effect on both participation equation, from 0.13 to 0.15, and level equation, from 2.21 to 2.25.

All transaction variables have statistically significant effects at least to some extent. The indicator variable for eurozone has a positive, statistically significant effect in all equations, while domestic credit per GDP in the destination economy has a significant effect only in specification (5). The effect of openness of the destination economy remains puzzling: Imports per GDP has a positive statistically significant effect and exports per GDP has a negative statistically significant effect on both equations and in both specifications.

Including trade and regional dummies has some effect on the results by somewhat reducing the size of the coefficients of the other variables. Bilateral trade itself has a statistically significant positive effect on the participation equation, but not on the level equation, again pointing out to a possible information effect.

All in all, the results for equity and debt security holdings are quite similar, with perhaps a bit surprisingly the information variables appearing to play a larger role with the debt security holdings. One might think that information is more crucial with the more risky equity investments, but the results may well point out to the problem of proxying information effects acknowledged in previous section: the employed variables may well be not the most appropriate ones in capturing the information effects crucial for cross-border investments.

Keeping in mind the limitations of the data and thus considering these results to be merely of illustrative nature, the double-hurdle appears to be working well in estimating the effects of determinants of cross-border investments. The results are qualitatively not too different from the more established approaches, so there is no imminent reason to suspect that the

de_{ij}	(4)		(5)		(6)	
	P	L	P	L	P	L
gdp_i	-0.13**** (0.01)	-0.85**** (0.08)	0.06 (0.06)	-0.49 (0.38)	0.08 (0.06)	-0.39 (0.39)
gdp_j	0.12**** (0.00)	0.80**** (0.05)	0.12**** (0.01)	0.79**** (0.06)	0.08**** (0.01)	0.69**** (0.11)
$dist_{ij}$	-0.06**** (0.01)	-1.03**** (0.09)	0.06**** (0.01)	-0.02 (0.09)	0.07**** (0.01)	0.06 (0.11)
$banksubs$	-	-	0.49**** (0.11)	-0.08 (0.35)	0.41**** (0.11)	-0.11 (0.35)
$lawscan$	-	-	1.46 (2817.20)	1.47 (4114.90)	1.28 (251.97)	1.58 (386.13)
eu	-	-	0.15*** (0.04)	2.25**** (0.26)	0.13*** (0.04)	2.21**** (0.27)
$domcred_j$	-	-	0.02* (0.01)	0.26** (0.10)	0.01 (0.01)	0.26 (0.11)
$euro$	-	-	0.21*** (0.08)	2.02**** (0.29)	0.18** (0.07)	2.09**** (0.29)
imp_j	-	-	0.17**** (0.03)	0.62** (0.30)	0.15**** (0.03)	0.57* (0.30)
exp_j	-	-	-0.10**** (0.02)	-0.68*** (0.24)	-0.09**** (0.02)	-0.71*** (0.24)
$trade$	-	-	-	-	0.04**** (0.01)	0.11 (0.09)
regional	No	No	No	No	Yes	Yes

Significance at the 10%, 5%, 1% and 0.1% levels is denoted by *, **, *** and **** respectively. Standard errors in parentheses.

Table 9: Marginal effects of the independent variables on cross-border debt security holdings in the double-hurdle model

	(1)*	(2)	(3)	(4)	(5)	(6)
N	947	1455	1455	1868	1455	1455
σ	2.43	1.95	1.94	2.54	1.55	1.54
	(0.10)	(0.06)	(0.06)	(0.10)	(0.05)	(0.05)
Covariance	0.69	0.32	0.37	1.68	0.45	0.45
	(0.31)	(0.18)	(0.19)	(0.23)	(0.14)	(0.14)

*This specification is estimated with data only on years 2004-2007 and 2013.

Standard errors in parentheses.

Table 10: Estimation results for the double-hurdle model of equity holdings

double-hurdle model is not working as expected. However, as the model is highly non-linear, it would benefit from being estimated with a larger and more diversified pool of data with more observations and more variation. All in all, as pointed out before, the double-hurdle model is able to detect the slight differences in the processes determining the dichotomous cross-border investment decision and the decision of how much to invest, which would not have been easy to detect if only the classic or the dichotomous approach were employed.

5 Conclusions and proposal for future research

In this thesis I study the determinants of international investments with the help of the gravity model of international financial asset trade. The relevant literature is discussed and a theoretical framework for gravity in cross-border investments is presented. I compare three different empirical approaches; the classic approach that studies the determinants of the observed levels of cross-border holdings by a fixed effects panel model, the dichotomous approach that studies the effects of determinants on the probability of there being a positive cross-border investment by a probit model and finally an approach which combines the two previous ones by a two-staged double-hurdle model. I propose that the correct method to estimate the determinants of cross-border investments could be the double-hurdle model, previously overlooked in estimating gravity models of financial asset trade.

International financial asset trade is characterized by two phenomena: home bias and a dominance of observed bilateral holdings equaling zero. The gravity model can be used to explain these phenomena by studying the determinants of cross-border investment holdings and flows. The classic approach of Portes and Rey (2005) to estimating the effect of these determinants is interested in the level of the cross-border investments observed, and although having been very successful in explaining the first phenomenon, has completely overlooked the second. An alternative and more appropriate approach is to use a model for limited dependent variables. One can treat the gravity equation as binary, and take interest in the dichotomous cross-border investment decision, as Drakos et al. (2014) do. This way also zero observations are addressed and allowed to yield information. However, the dichotomous approach loses valuable information by not considering the levels of the observed holdings. A synthesis of these two approaches is needed.

The double-hurdle mode of Cragg (1971) is apt for the task. The double-hurdle is a model for limited dependent variables, but in addition it has a specific feature: It allows two decisions, a participation decision and a level decision, and it allows them to be determined by to separate processes. The

cross-border investment decision is intuitively easy to deconstruct into two distinct decisions. An investor considering investing in a given foreign country must first make up his mind on the dichotomous question of investing or not. If, and only if, the decision is positive, he then decides on the level of his investment. Thus, there can be thought to be a participation decision and a level decision that are distinct from each other. There is no reason why the decision to invest should be determined by a similar process as the decision of how much to invest. Thus the double-hurdle is a correct empirical approach to estimating gravity equations of international financial asset holdings and flows.

I illustrate my argument with an empirical application to the external assets of Finnish residents. Indeed I find that the decision to invest and the decision of how much to invest are different for Finnish investors: The classic approach to estimating the gravity equations and the dichotomous approach have qualitatively slightly different results, suggesting that the two decisions may be determined by different processes. When estimated with the double-hurdle, differences become clearer: the magnitudes of the effects, their statistical significance and in certain cases even the signs of the effects are different between the participation and the level equation.

Even though these results are merely illustrative and surely have many caveats, they do point out that the double-hurdle model is well worth future research in the context of cross-border investments. A natural next step would be to estimate a double-hurdle gravity model for a pool of observations. The set-up of this thesis could be extended in various manners. I have considered annual data, but data on external assets on Finnish investors is also available on a quarterly and monthly frequency. Furthermore, the model could be disaggregated over different sectors, and for example the determinants of sophisticated institutional investor and households could be compared. Increasing the set of investor countries would also help find more solid results. The CPIS data set allows extending the study to cover all 242 countries that report their external portfolio assets to IMF.

It would also be interesting to account for the serial correlation in holdings. This could be done by introducing a dynamic version of the double-hurdle model or in a more simple manner by taking the difference in stocks of holdings instead of holdings as the dependent variable. In addition to cross-border portfolio investments there are a lot of other possible applications for the double-hurdle model to be found in the financial markets. This model that has been previously overlooked in the field of finance, investments and financial markets, should be put to more use. The double-hurdle model would surely prove to be a fertile ground for future research.

6 References

Anderson, J.E., 2010. The Gravity Model. NBER Working Paper 16576.

Anderson, J.E., 1979. A Theoretical Foundation for the Gravity Equation. *The American Economic Review* vol. 69, pp. 106-116.

Anderson, J.E., van Wincoop, E., 2004. Trade costs. *Journal of Economic Literature* vol. 42, pp. 691-751.

Anderson, C.W., Fedenia, M., Hirschey, M., Skiba, H., 2011. Cultural Influences on Home Bias and International Diversification by Institutional Investors. *Journal of Banking & Finance* vol. 35, pp. 916-934.

Aristei, D., 2013. A Blinder-Oaxaca Decomposition for Double-Hurdle Models with an Application to Migrants' Remittance Behaviour. *Applied Economics Letters* vol. 20, pp. 1665-1672.

Aviat, A., Coeurdacier, N., 2007. The Geograhpy of Trade in Goods and Asset Holdings. *Journal of International Economics* vol. 71, pp. 22-51.

Balli, F., Balli, H.O., Basher, S.A., 2013. The Determinants of Bilateral Cross-Border Bond and Equity Flows: a Sectoral Analysis. Available at www.syedbasher.org/wp/sector_portfolio.pdf.

Berkel, B., 2007. Institutional Determinants of International Equity Portfolios - a Country-Level Analysis. *The B.E. Journal of Macroeconomics* vol. 7, article 34.

Berkel, B., 2006. European Financial Market Integration and German Cross Border Portfolio Flows. EFA 2006 Zurich Meetings. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=903705.

Blundell, R., Meghir, C., 1987. Bivariate Alternatives to the Tobit Model. *Journal of Econometrics* vol. 34, pp. 179-200.

Blundell, R., Ham, J., Meghir, C., 1987. Unemployment and Female Labor Supply. *The Economic Journal* vol.97, supplement: conference papers, pp. 44-64.

Box, G.E.P., Cox, D.R., 1964. An Analysis of Transformations. Journal of the Royal Statistical Society. Series B (Methodological) vol. 26, pp. 211-252.

Brennan, M.J., Cao, H.H., 1997. International Portfolio Investment Flows. The Journal of Finance vol. 52, pp. 1851-1880.

CEPII. GeoDist database. Available at http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=6.

Coeurdacier, N., Martin, P., 2009. The Geography of Asset Trade and the Euro: Insiders and Outsiders. Journal of the Japanese and International Economics vol. 23, pp. 90-113.

Coeurdacier, N., Guibaud, S., 2011. International Portfolio Diversification Is Better Than You Think. Journal of International Money and Finance vol. 30, pp. 289-308.

Coeurdacier, N., Kollman, R., Martin, P., 2010. International Portfolios, Capital Accumulation and Foreign Assets Dynamics. Journal of International Economics vol. 80, pp. 100-112.

Coval, J. D., Moskowitz, T. J., 1999. Home Bias at Home: Local Equity Preference in Domestic Portfolios. The Journal of Finance vol. 54, pp. 2045-2073.

Cragg, J. G., 1971. Some Statistical Models for Limited Dependent Variables with Application to the Demand of Durable Goods. Econometrica vol. 39, pp. 829-844.

Dahlqvist, M., Pinkowitz, L., Stulz, R., Williamson, R., 2003. Corporate Governance and the Home Bias. Journal of Financial and Quantitative Analysis vol. 38, pp. 87-110.

Drakos, K., Kyriazidou, E., Polycarpou, I., 2014. A Dynamic Gravity Model for Global Bilateral Investment Holdings. International Symposia in Economic Theory and Econometrics vol. 23, pp. 126-152.

Felettigh, A., Monti, P., 2008. How to Interpret the CPIS Data on the Distribution of Foreign Portfolio Assets in the Presence of Sizable Cross-Border Positions in Mutual Funds. Evidence for Italy and the Main Euro-area Countries. *Questioni di Economia e Finanza, Banca d'Italia*.

French, K., Poterba, J., 1991. Investor Diversification and International Equity Markets. *American Economic Review* vol. 81, pp. 222-226.

Garcia, B., 2013. Implementation of a Double-hurdle Model. *The Stata Journal* vol. 13, 776-794.

di Giovanni, J., 2005. What Drives Capital Flows? The Case of Cross-border M&A Activity and Financial Deepening. *Journal of International Economics* vol. 65, pp. 127-149.

Goetzmann, W.N., Kumar, A., 2008. Equity Portfolio Diversification. *Review of Finance* vol. 12, pp. 433-463.

Grinblatt, M., Keloharju, M., 2000. The Investment Behaviour and Performance of Various Investor Types: a Study of Finland's Unique Data Set. *Journal of Financial Economics* vol. 55, pp. 43-67.

Head, K., Mayer, T., 2013. Gravity Equations. Workhorse, Toolkit and Cookbook. Available at http://strategy.sauder.ubc.ca/head/papers/headmayer_revised.pdf.

Heckmann, J.J., 1979. Sample Selection Bias as a Specification Error. *Econometrica* vol. 47, pp. 153-161.

Heckmann, J.J., 1976. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator of Such Models. *Annals of Economic and Social Measurement* vol. 5, pp. 475-492.

Holloway, G., Barret, C.B., Ehui, S., 2005. Bayesian Estimation of the Double Hurdle Model in the Presence of Fixed Costs. *Journal of International Agricultural Trade and Development* vol. 1, pp. 17-28.

International Monetary Fund. Coordinated Portfolio Investment Survey, Data and Metadata. Available at <http://cpis.imf.org/>.

Jones, A.M., 1989. A Double-hurdle Model of Cigarette Consumption. *Journal of Applied Econometrics* vol. 4, pp. 23-39.

Jones, A.M., Yen, S.T., 2000. A Box-Cox Double-hurdle Model. *The Manchester School* vol 68, pp. 203-221.

Lane, P.R., Milesi-Ferretti, G.M., 2005. The International Equity Holdings of Euro Area Investors. IIS Discussion Paper No. 104/December 2005.

Lane, P.R., Milesi-Ferretti, G.M., 2008. International Investment Patterns. *The Review of Economics and Statistics* vol. 90, pp. 538-549.

Li, S., Faruquee, H., Yan, I.K., 2004. The Determinants of International Portfolio Holdings and Home Bias. IMF working paper. Available at www.imf.org/external/pubs/ft/wp/2004/wp0434.pdf.

Martin, P., Rey, H., 2004. Financial Super-markets: Size Matters for Asset Trade. *Journal of International Economics* vol. 64, pp. 335-361. First published in 1999 as a CEPR discussion paper 2232, available at www.cepr.org/active/publications/discussion_papers/dp.php?dpno=2232.

Mayer, T., Zignago, S., 2011. Notes on CEPII's Distances Measure: the GeoDist Database. CEPII Working paper 2011-25. Available at <http://www.cepii.fr/CEPII/en/publications/wp/abstract.asp?NoDoc=3877>.

Newman, C., Henchion, M., Matthews, A., 2003. A Double-hurdle Model of Irish Household Expenditure on Prepared Meals. *Applied Economics* vol. 35, pp. 1053-1061.

Obstfeld, M., Rogoff, K., 2000. The Six Major Puzzles in International Macroeconomics: Is There a Common Cause? Published in *NBER Macroeconomics Annual 2000*, volume 15. MIT Press.

Obstfeld, M., Rogoff, K., 1999. *Foundations of International Macroeconomics*. Massachusetts Institute of Technology.

Okawa, Y., van Wincoop, E., 2012. Gravity in International Finance. *Journal of International Economics* vol. 87, pp. 205-215.

Portes, R., Rey, H., 2005. The Determinants of Cross-border Equity Flows. *Journal of International Economics* vol. 65, pp. 269-296.

First published in 1999 as a CEPR discussion paper 2225, available at www.cepr.org/active/publications/discussion_papers/dp.php?dpno=2225.

Portes, R., Rey, H., 1998. The Euro and International Equity Flows. *Journal of the Japanese and International Economies* vol. 12, pp. 406-423.

Portes, R., Rey, H., Oh, Y., 2001. Information and Capital Flows. The determinants of Transactions in Financial Assets. *European Economic Review* vol. 45, pp. 783-796.

Salo, A., 2014. Maksutaseen tunnistamattomat ja virheelliset erat. Bank of Finland, BoF-online 4/2014. Available in Finnish at www.suomenpankki.fi/fi/julkaisut/selvitykset_ja_raportit/bof_online and as an English abstract at www.suomenpankki.fi/en/julkaisut/selvitykset_ja_raportit/bof_online.

Sercu, P., Vanpee, R., 2007. Home Bias in International Equity Portfolios: a Review. Katholieke Universiteit Leuven, Faculty of Economics and Applied Economics. Available at <https://lirias.kuleuven.be/bitstream/123456789/175483/1/AFL0710.pdf>.

Sorensen, B.E., Wu, Y., Yosha, O., Zhu, Y., 2007. Home Bias and International Risk Sharing: Twin Puzzles Separated at Birth. *Journal of Money and Finance* vol. 26, pp. 587-605.

Zorzi, R., 2009. Data on Taxation of Savings Income in the Context of BOP/IIP Compilation. Background Material for the WG-ES Thematic Meeting in Warsaw 2009, Oesterreichische Nationalbank.

Zucman, G., 2013. The Missing Wealth of Nations: Are Europe and the U.S. Net Debtors or Net Creditors? *The Quarterly Journal of Economics* vol. 128, pp. 1321-1364.