The Determinants of International Equity Holdings: Information vs. Culture

Erkin Diyarbakirlioglu [*]

Sorbonne Business School
21 rue Broca, 75005 Paris – France
&
Galatasaray University, Faculty of Economic and Administrative Sciences
Çırağan cad. n° 36, 34457 Ortakoy / Istanbul – Turkey

Abstract

This paper examines whether the effect of physical distance on the geography of international portfolio investments undergoes an informational or a cultural channel. Using two measures to disentangle the effects of the physical distance, namely the *cultural distance* (Kogut and Singh, 1988), and a new measure of the *information distance*, I analyze a panel dataset on a sample of 24 source countries' bilateral equity assets held in various destination countries for end-2006. Regressions suggest that the observed geographical patterns of bilateral portfolio investments can be explained by information asymmetries rather than cultural affinities between countries.

Key words: information distance; cultural distance; gravity model; home bias

EFM classification codes: 370, 630

[*] E-mail: Erkin.Diyarbakirlioglu@malix.univ-paris1.fr, Tel.: +33 1 53552704, Fax: +33 1 53552703.

1 Introduction

Recent literature on home bias and international portfolio diversification suggests the following results regarding the geographical patterns of aggregate and bilateral asset holdings:

1) Despite the frequently-claimed gains from global portfolio diversification and the easier access to financial markets worldwide (Baele and Inghelbrecht, 2009; Chiou, 2009), home bias remains still high across investors (Sorensen *et al.*, 2007; Baele *et al.*, 2007).
2) The propensity to invest cross-border declines with the home country's economic mass, implying that the domestic bias is particularly stronger among emerging-market countries although they are likely to benefit more from global portfolio diversification (Campbell and Kraussl, 2007; Driessen and Laeven, 2007).
3) Investors concentrate their already trivial cross-border assets in a handful of destinations (Hau and Rey, 2008) which generally tend to be mature and developed countries, and the huge volumes of cross-border capital flows mainly turn around the industrialized world in contrast to the predictions of the benchmark neoclassical model with frictionless markets (Papaioannou, 2009; Prasad *et al.*, 2007).
4) The geography of bilateral portfolio holdings shows evidence of a preference toward physically proximate alternatives, so that distance and stock market correlations reveal out to be significantly positive correlates of the bilateral holdings (Lane and Milesi-Ferretti, 2008; Aviat and Coeurdacier, 2007; Portes and Rey, 2005).

These observations make it hard to justify an investor portrait whose objective is simply to maximize the expected risk-adjusted-return on his investments. Pioneering studies in the field such as Adler and Dumas (1983) and Solnik (1974), propose that agents should be better off investing into an identical risky portfolio partly hedged against the exchange rate risk (under the assumption of deviations from the purchasing power parity), in which every asset is represented in proportion to its relative weight in the world-market portfolio. However, these models along the lines of rational decision-making and market efficiency paradigms commonly fail to replicate the size and the asymmetries of the domestic portfolio bias as well as the geographical

underpinnings of bilateral asset holdings. For example, using a sample of 25 countries, Baele *et al.* (2007) report that local investors exhibit an average home bias between 0.7 and 0.8. Sorensen *et al.* (2007) show that the average home bias across OECD countries is about 67% as of the end-2003. Moreover, their estimates are far from being homogenous in that emerging countries exhibit substantially higher home bias than developed countries.[1]

Studies on the reasons as to why investors overweight the home country in their portfolio allocations mainly fall into two broad categories (French and Poterba, 1991). Among the so-called *institutional explanations*, a number of attempts tackled the issue in terms of direct barriers to international investments such as capital controls (Errunza and Losq, 1985; Stulz, 1981; Black, 1974) or other forms of market imperfections such as transaction costs (Rowland, 1999). However, both the gradual liberalization of capital markets starting from the mid-1980's and the relatively higher volumes of cross-border trading as revealed by high turnover rates in international transactions rule out these explanations (Warnock, 2002; Tesar and Werner, 1995). A second group of studies highlight investors' hedging motives against domestic price uncertainty due to i) deviations from purchasing power parity (Adler and Dumas, 1983), ii) non-traded consumption goods (Stockman and Dellas, 1989; Cooper and Kaplanis, 1994), or iii) non-tradable factors such as human capital (Baxter and Jermann, 1997).[2] Finally, the home bias can also be due to information asymmetries between local and foreign investors. Brennan and Cao (1997) and Gehrig (1993) provide theoretical models where home bias arises from an informational advantage possessed by local investors on their home market over foreign investors. Recently, Barron and Ni (2008) and Ni (2009) add up to this literature by showing that, in addition to the asymmetric information, portfolio size also contributes to the likelihood of investors to be home biased. Empirical evidence with respect to the impact of the asymmetric information on home

[1] For example, the average home bias across four emerging markets studied by Baele *et al.* (2007) is about 96%.

[2] Cooper and Kaplanis (1994) note that inflation hedging is an unlikely reason of home bias for reasonable degrees of investors' risk aversion. Baxter and Jermann (1997) show that investors should even sell short their own market to hedge human capital risk because of the strong correlation prevailing between returns on physical capital and human capital. More recently, using Swedish household portfolio data, Massa and Simonov (2006) show that investors do not engage in hedging nonfinancial and financial income, but instead they tilt their portfolio toward stocks geographically and professionally close to them.

bias is, however, somewhat mixed.[3]

Other studies account for the effect of investor psychology on the portfolio selection problem. Putting forward several theories and concepts drawn from the behavioral finance school of thought, the central premise is that individuals are only quasi-rational in their decision-making process (Ricciardi, 2008a) and act mainly according to the principles of Kahneman and Tversky's (1979) prospect theory. The main insight is to reject the fully rational model of individual decision-making by introducing a series of cognitive and affective aspects likely to influence an investors' risk perception in uncertain or risky decision-making contexts.[4] In applied work related to domestic or international asset allocation, the literature has addressed a number of concepts like the "familiarity bias" (Chan *et al.*, 2005; Huberman, 2001; Grinblatt and Keloharju, 2001), "narrow framing" (Magi, 2009; Nocetti, 2006), or investors' relative optimism and perceived competence about domestic stocks (Suh, 2005; Strong and Xu, 2003). In short, although it is acknowledged that direct barriers to international investments are nowadays an unlikely reason to explain the home bias, the literature still lacks a fully convincing case, whereby the ongoing debate mainly contrasts the informational vs. behavioral explanations of the home bias puzzle, none of them excluding each other.

A common feature among these attempts is that they all investigate the extent of the aggregate or the "country level" home bias (Hau and Rey, 2008; Karlsson and Norden, 2007), i.e. items number 1 and 2 stated above. In this study, following Lane and Milesi-Ferretti (2008) and Portes and Rey (2005), I consider the aggregate home bias as given and focus on the determinants of bilateral holdings instead. Specifically, I examine the geography of cross-border equity investments to shed light on the "distance" (Portes and Rey, 2005; Aviat and Coeurdacier, 2007) puzzle mentioned in items 3 and 4 above. The underlying theoretical assumption is that, in a fully

[3] Dvorak (2005) finds that foreign institutions in Indonesian stock market enjoy better information due to their experience and expertise. Lütje and Menkhoff (2007) report survey results that local investors fail to materialize the informational advantage they pretend to possess. On the opposite, Malloy (2005) and Hau (2001) argue that local analysts and investors outperform foreigners, a finding that supports the link between physical proximity and information quality.

[4] Ricciardi (2008a, 2008b) provides two extensive and up-to-date surveys of the behavioral finance theory in comparison to the traditional finance theories. The author presents a formal introduction to theories related to risk perception and other concepts that influence the individual's decision-making process under conditions of risk and uncertainty. For other discussions, see also Barberis and Thaler (2003) or Shleifer (2000).

integrated global economy with frictionless goods and asset markets, 1) investors should hold identical portfolios (Lane and Milesi-Ferretti, 2008) and 2) capital flows from rich to poor countries where marginal returns are higher (Papaionnaou, 2009; Lucas, 1990). Yet, to the extent that investors concentrate their foreign portfolio holdings within a handful of proximate and highly correlated destinations, it is unlikely to justify any risk-sharing motive within the existing bilateral portfolio holdings worldwide.

I argue that the preference revealed by investors toward geographically proximate alternatives would undergo two different but somewhat complementary viewpoints. On the empirical side, geography is documented as a key determinant of the investment decisions either domestically (Huberman, 2001; Coval and Moskowitz, 1999) or internationally (Lane and Milesi-Ferretti, 2008; Chan *et al.*, 2005; Sarkissian and Schill, 2004). Intuitively, however, distance should not capture the costs of trading in securities since, unlike physical goods, assets are *weightless*. Distant stocks (or countries) could even be thought of as being better hedge instruments against local or regional risks prevailing among countries geographically closer to each other due to the stronger political, economic, and trade linkage they typically share. So, why this is not the case and where does this obvious effect of the distance comes from? According to Portes and Rey (2005, p. 270), "the most natural explanation is that informational frictions are positively correlated with distance" which "is a barrier to interaction among economic agents and […] to cultural exchange". From this perspective, distance is mostly related to an asymmetric information framework, implying that "investors buy […] securities about which they have enough information" (Merton, 1987) so that they prefer closer geographies in forming their portfolios.

Besides, the distance puzzle could also be related to some kind of *familiarity bias* underlying the investor behavior. Ricciardi (2008a) defines the familiarity bias simply as *an inclination that alters individuals' risk perception* (p. 101). The concept conjectures that individuals feel more comfortable with risks they feel familiar with, so that they use "heuristic simplifications in their decision-making process" (Massa and Simonov, 2006). Hence, geography would help to capture such mental shortcuts in the context of portfolio selection where investors typically prefer to

allocate across nearby stocks or markets. Although it is by now a well-known fact that investors prefer familiar stocks (Huberman, 2001) or markets located nearby (Aviat and Coeurdacier, 2007; Grinblatt and Keloharju, 2001), whether the puzzling effect of the geography goes through the asymmetric information or the familiarity-bias framework, remains unclear. Massa and Simonov (2006) note that *familiarity-driven investment is a rational response to information constraints as opposed to a behavioral heuristic* (p. 634). Contrasting the information-driven with the behavioral hypotheses, the authors conclude that 1) familiarity mostly affects less informed investors and 2) the more sophisticated the investor is, the weaker is the effect of behavioral familiarity on decision-making. A similar result has also been provided by Grinblatt and Keloharju (2001) who note that *the influence of distance and culture on stockholdings and trades is smaller* […] *for more sophisticated household investors* (p. 1054). To summarize, prior studies show that geographical distance is frequently used in empirical work on home bias and the patterns of international investments, and evidence suggest that "distance comes up remarkably strongly with the predicted (negative) sign" (Portes *et al.*, 2001). However, it is so far unclear whether the intriguing effect of the distance on the individual's decision making and portfolio selection goes through an asymmetric information or a behavioral channel. As noted by Portes and Rey (2005), distinguishing purely behavioral familiarity effects from those originated by informational asymmetries remains a challenge to the existing literature.

I propose to contribute to this body of research by analyzing the determinants of bilateral equity holdings within a micro-founded model of international capital market equilibrium. Setting the bilateral stock of equity holdings as the dependent, I estimate a gravity model for international portfolio assets using not only the physical distance but also the *informational* and *cultural* distances between countries. Specifically, I construct two artificial variables likely to proxy for informational and cultural frictions across countries, and also likely to underlie the puzzling effect of the geographical distance. Namely, I propose a measure of the *information distance* between two countries A and B by plugging data on bilateral phone call traffic into a

simple distance formula developed by Cilibrasi and Vitanyi (2007).[5] In turn, I define the *cultural distance* between A and B using a formula originally proposed by Kogut and Singh (1988) and numerical scores on four country-specific cultural dimensions identified by Hofstede's work on cultural affinities.[6] The main motivation to include these two measures essentially follows Massa and Simonov's (2006), and Portes and Rey's (2005) inquiries on the competing hypothesis of behavioral-based vs. information-based familiarity channel of the geography. While investigating the relative effects of geographical, information and cultural distance on bilateral equity investments, I also control for a number of other gravity-type variables mainly motivated from the previous literature as well. Namely, I expand the scope of the analysis by including an array of additional variables classified into the following categories: 1) Economic development, 2) Openness, 3) Familiarity, 4) Transparency, and 5) Portfolio Diversification.

Data limitations regarding the estimation of bilateral or aggregate investment stocks have usually meant important concerns for the related literature,[7] restricting the scope of the analysis into a single country (for instance, Sweden in Karlsson and Norden, 2007; Japan in Kang and Stulz, 1997; Australia in Mishra and Daly, 2006; or the United States in Ahearne *et al.*, 2004; and Dahlquist *et al.*, 2003). Other studies have directly focused on the patterns of foreign direct investment (Stein and Daude, 2007; Wei, 2000) or those of the international banking assets (Aviat and Coeurdacier, 2007), for which comparable data on investment positions is readily available.[8] In this study, I use survey data as of the end-2006 from the International Monetary Fund to analyze the foreign portfolio allocations over a sample of 24 countries. Since the first time

[5] To my knowledge, this paper is the first attempt to propose such a direct quantitative measure of information distance in finance.

[6] See Hofstede (1983). Details regarding the calculations of these two distance measures are provided below.

[7] Among studies taking into account a broader perspective, cross-border holdings were commonly estimated using flow data from balance of payments statistics (see, for instance, Baele *et al.*,2007; Portes and Rey, 2005; Bekaert and Harvey, 2000). However, as warned by Cleaver and Warnock (2003), the use of flows data can be misleading when one looks for obtaining stock positions due to high turnover rates observed in international capital flows and the very nature of the balance of payments methodology. For example, Warnock (2002) points out that flow data does not allow concluding on the true origin of foreign investment since the host country and the intermediary country from which the operation is held are not necessarily the same.

[8] See also Chan *et al.* (2005) and Gelos and Wei (2005) who employ data on mutual fund's cross-border equity allocations. For example, Gelos and Wei (2005) use data on 137 equity funds' country-level portfolio allocation (managing a total by US$ 44 billion of assets in emerging markets), while Chan *et al.* (2005) use of mutual fund equity holdings from 26 developed and developing countries with a breakdown across 48 destinations.

it had been published in 1997, the *Coordinated Portfolio Investment Survey* (CPIS) is being released on an annual basis by the IMF from 2001 onwards. It has the main advantage of directly providing holdings data on equities, as well as short and long-term bonds. Although more than 70 reporting countries had participated to the 2006 survey, the sum of the total foreign equity assets held by this 24-countries sample is quite representative and covers about 72% of the total assets reported.[9]

In line with previous work such as Lane and Milesi-Ferretti (2008) or Aviat and Coeurdacier (2007), I use a gravity model as the basis of the empirical investigations. In its well-known original setup in physics, the model postulates an equilibrium relationship between two distinct objects' masses and the physical distance between them. Since the pioneering work by Tinbergen (1962), the gravity model constitutes an important toolbox to trade economists, while its implementation into the financial globe is relatively recent. Using capital flow data, studies such as Papaioannou (2009) or Portes and Rey (2005) have shown that the gravity model can also explain the patterns of international trade in securities as good as it does in physical goods. Being well supported on the empirical ground, a theoretical support is provided by Martin and Rey (2004) who developed a two-country equilibrium model from which a gravity-style relationship emerges naturally. In short, the gravity model postulates that the volume of bilateral transactions (or holdings) between two countries is an increasing function of their respective economic masses and a decreasing function of the trading cost among them. In applied work, the economic size and the trading cost are generally substituted by countries' respective GDPs and the physical distance respectively.

The paper is organized as follows. The next section reviews the theoretical framework proposed by Martin and Rey (2004), which also serves as the basis of the subsequent econometric analysis. The third section introduces the methodology and the data set. The fourth section discusses the results of the estimations. The last section concludes.

[9] For others papers making use of the CPIS data, see, for example, Lane and Milesi-Ferretti (2008), Mishra and Daly (2006) or Faruqee *et al.* (2004).

2 A Model of International Asset Holdings

In this section, I provide a brief overview of the theoretical setting that I will make use of throughout the empirical analysis. The framework follows the general equilibrium model under incomplete markets introduced in a two-country setting by Martin and Rey (2004), and also studied by Aviat and Coeurdacier (2007) and Faruqee *et al.* (2004).[10] The objective is to lay out a review of the benchmark theory upon which I base the econometric specification of the gravity equation used throughout the empirical analysis.

Let us assume an international capital market where countries are indexed by i or j. Each country is populated by n_i agents with intertemporal, concave and strictly decreasing utility functions.[11] At time t each agent (i.e. investor) $h_i \in \{1, \dots, n_i\}$ is endowed y units of tradable good (i.e. the numéraire) which he can either consume or invest in a set of risky projects.[12] The total number of the risky projects developed by an agent is z_{hi} and the cost of developing a new project is a differentiable function $f(z_{hi})$ with $f'(z_{hi}) > 0$ and $f''(z_{hi}) > 0$. Thus, not only the cost of developing a new project is an increasing function of z_{hi} but the associated marginal cost is also increasing with the number of projects already developed.

The next period $t + 1$ comprises S different states of the world, each with equal probability $Pr(i = S) = 1/S$. Following this setup, each risky project can be considered as an Arrow-Debreu style security because its payoff next period is either equal to d if state $i \in \{1, \dots, S\}$ occurs, and 0 otherwise. There are no intermediary income streams such as labor income, so that the dividends d, are the unique source of consumption next period. Different securities serve as hedge instruments against different sources of risk; consequently, each security is an imperfect substitute of another one. This feature of the model simply implies that diversification across

[10] In an earlier version of their 2008 paper, Lane and Milesi-Ferretti (2004) distinguish three alternatives for modeling international portfolio holdings. The first approach, due to Obstfeld and Rogoff (2000), is a static model in which frictions in good markets lead to domestically biased portfolios even tough financial markets are complete. The *N*-country generalization of the Obstfeld-Rogoff study is provided by Lane and Milesi-Ferretti (2004). The second approach by Davis *et al.* (2000) assumes a dynamic model of portfolio allocation and consumption under incomplete markets. The third approach (which is also the one we adopt here), comes from Martin and Rey (2004) who develop a model of bilateral asset holdings from which a gravity-type model emerges naturally.

[11] The concavity assumption implies also that agents are risk-averse.

[12] Intuitively, the term "agents" substitutes investors and "risky projects" substitutes a risky financial instrument, such as equities for example.

securities is beneficial. Nevertheless, the number of the states of the world is bigger than the total number of Arrow-Debreu securities, implying that 1) the market is incomplete, and 2) at the limit, diversification cannot eliminate all the risk an agent bears.

Risky securities developed by agents in different countries are traded on a frictional international capital market. The residents of the country i pay a transaction cost $\tau_i^j > 0$ when they trade overseas or earn a dividend gain in a foreign country $j \neq i$. Letting p_{h_j} be the price of a share of the project developed by agent h_j and s_{hi}^j be the demand of $h_i \neq h_j$ for an asset traded in country j, the amount paid by h_i to purchase one share of such an asset is given by $p_{h_j} s_{hi}^j (1 + \tau_i^j)$ if the asset pays a dividend next period, and the agent h_i i.e. the holder of the asset, receives $d(1 - \tau_i^j)$ per share of project he purchased. Assuming that the transaction costs apply to the buyer of the asset, the budget constraint of the representative agent in country i can be expressed as follows,

$$y_i + \sum_{i \in \{1,\dots,z_{hi}\}} \alpha_{hi}^i p_{hi}^i = c_{1,hi} + f(z_{hi}) + \sum_{\substack{i \in \{1,\dots,n_i\} \\ i \neq h_i}} p_i s_{hi}^i + \sum_{j \in \{1,\dots,n_j\}} p_j s_{hi}^j (1 + \tau_i^j) \quad (1)$$

On the left-hand side of the equation, beside the initial endowment y_i, the investor sells a portion α_{hi}^i of the securities he developed himself. Put another way, the coefficient alpha represents the investor's diversification level, and by construction, the term $(1 - \alpha_{hi}^i)$ corresponds to the share of projects developed by h_i but which does not float on the market. On the right-hand side, we observe the agent's consumption in the first period, $c_{1,hi}$, the cost he bears to develop new projects, $f(z_{hi})$, and his demand for assets developed by agents other than himself in country i and in country j on which a transaction cost $(1 + \tau_i^j)$ already applies.

Following the standard rational expectations model of portfolio choice[13], each agent maximizes a time-additive utility of the form,

[13] See Uppal and Wang (2003), p. 2467.

$$E[U_{hi}] = c_{1,hi} + \delta \cdot E\left[\frac{\left(c_{2,hi}\right)^{(1-1/\sigma)}}{1-1/\sigma}\right] \quad (2)$$

where δ is the subjective discount rate of the next period utility from consumption and corresponds to the inverse of investor's coefficient of risk aversion. Obviously, the latter one is different from zero, which also captures the feature that all agents are risk-averse. According to the payoff structure imposed to Arrow-Debreu securities and the hypothesis that all states of the world next period have equal probability, we can explicitly write the expected utility as,

$$E[U_{hi}] = c_{1,hi} + \delta \cdot E\left[\frac{1}{S}\frac{1}{1-\frac{1}{\sigma}}\left(\sum_{i\in z_{hi}}\left((1-\alpha_{hi}^{i})d\right)^{\left(1-\frac{1}{\sigma}\right)} + \sum_{\substack{i\in n_i \\ i\neq h_i}}\left(ds_{hi}^{i}\right)^{\left(1-\frac{1}{\sigma}\right)} + \sum_{j\in n_j}\left((1-\tau_i^j)s_{hi}^{j}\right)^{\left(1-\frac{1}{\sigma}\right)}\right)\right] \quad (3)$$

Taking the expectation on the right-hand side and rearranging, we obtain,

$$E[U_{hi}] = c_{1,hi} + \frac{\delta}{S}\frac{d^{\left(1-\frac{1}{\sigma}\right)}}{1-\frac{1}{\sigma}}\left(\sum_{i\in z_{hi}}\left(1-\alpha_{hi}^{i}\right)^{\left(1-\frac{1}{\sigma}\right)} + \sum_{\substack{i\in n_i \\ i\neq h_i}}\left(s_{hi}^{i}\right)^{\left(1-\frac{1}{\sigma}\right)} + \left(1-\tau_i^j\right)\sum_{j\in n_j}\left(s_{hi}^{j}\right)^{\left(1-\frac{1}{\sigma}\right)}\right) \quad (4)$$

under the budget constraint given above. We observe that there are three elements making up the expected consumption in the second period. The first term in the parenthesis on the right-hand side of the equation is the expected utility from consumption due to payoffs of the projects developed by the agent himself and which does not float on the market. The second term corresponds to the utility from consumption the representative agent would obtain by dividend-paying projects developed by other agents in the local market. Finally, the third element corresponds to the expected utility from consumption backed by the dividend yields of the projects purchased abroad. Choosing the consumption in the first period, the number of risky projects to develop, the demand for domestic and foreign assets developed by other agents and the number of projects to keep in the second period, the representative maximizes the utility (4) subject to the budget constraint (1). From the first-order conditions, we obtain the following individual asset demands for projects traded in home and foreign countries respectively,

(5) $$s_{hi}^{i} = \left(\frac{\delta}{S}\right)^{\sigma} d^{\sigma-1} p_i^{-\sigma} \quad \text{and} \quad s_{hi}^{j} = \left(\frac{\delta}{S}\right)^{\sigma} d^{\sigma-1} p_j^{-\sigma} \frac{\left(1-\tau_i^j\right)^{\sigma-1}}{\left(1+\tau_i^j\right)^{\sigma}}$$

The market clears if the demand from both domestic and foreign agents for a project developed and traded in country j is equal to the supply, or more formally if $\sum_i n_i s_{hi}^j = 1, \forall j$, assuming that the number of shares for each project is normalized to one. Then, from the individual asset demand, we deduce the aggregate cross-border holdings of country i in country j as follows,

(6) $$Y_i^j = n_i \left(\left(\frac{\delta}{S}\right)^{\sigma} d^{\sigma-1} p_j^{-\sigma} \frac{\left(1-\tau_i^j\right)^{\sigma-1}}{\left(1+\tau_i^j\right)^{\sigma}} \right) n_j p_j$$

Rearranging, the expression above yields to the following log-linear relationship:

(7) $$\log\left(Y_i^j\right) = \sigma \log\left(\frac{\delta}{S}\right) + \log\left(n_i n_j\right) + (\sigma - 1)\log\left(r_j\right) - \log\left(\theta_i^j\right)$$

Equation (7) shows a theoretical gravity-style relationship on bilateral portfolio holdings held by the source country i in the host country j. The first term is a constant. The second term corresponds to the investor and host countries' market sizes. The third term reflects a return-chasing component with $r_j = d/p_j$. Finally, the last term stands for the international transaction costs where θ_i^j substitutes $\left(1-\tau_i^j\right)^{\sigma-1}/\left(1+\tau_i^j\right)^{\sigma}$. As noted by Aviat and Coeurdacier (2007), the key point is to determine the transaction costs which characterize the international capital market with frictions. Empirical studies typically employ the bilateral distance between the source and the destination countries as an accurate proxy of the transaction costs. However, as discussed in the introduction, whether the impact of the geography undergoes an informational or a behavioral channel remains an unanswered yet relevant question for existing studies.

3 Methodology and Data

3.1 Econometric Specification

In light of the theoretical detour above, I propose as the basis of the econometric analysis the

following log-linear model for portfolio assets held by investors in country i in the destination country j,

$$\log(Y_{ij}) = c + \alpha \log(M_i M_j) + \beta \log(\theta_{ij}) + \gamma \log(r_j) \quad (8)$$

with $\alpha > 0$ and $\beta < 0$. The product $M_i M_j$ stands for the size effect where M_i and M_j denote the respective economic masses of the source and destination countries. The parameter θ_{ij} stands for the cost of trading in international capital market. In this basic form, the equation represents a log-linearized structure of the original gravity model which had been frequently used to explain patters of international trade in goods. As emphasized previously, recent studies such as Portes and Rey (2005) or Portes *et al.* (2001) have shown that the model performs equally well when we consider trade in financial securities too. The log-log specification implies that parameter estimates are in terms of elasticities, i.e. the ratios of percentage changes in the corresponding variables. Due to the signs imposed to coefficients, the model considers the bilateral trade in securities between a pair of source and destination countries as an increasing function of their respective economic masses and a decreasing function of the trading cost between them.

When going to data, I substitute Y_{ij} by $ASSETS_{ij}$, i.e. the stock of equity assets held by country i (i.e. the source) in country j (i.e. destination), and the terms M_i and M_j by the countries respective GDPs. The term r_j captures the return-chasing component as shown in equation (7). I substitute it by the average return on the country j's broad stock market index over the period 2002:01-2006:12 (see below for further details). In equation (8), the key parameter of interest is β through which I capture the effect of international market frictions, θ_{ij}. The first proxy is the geographical distance between the source and host countries. I also control for two alternative distance concepts as mentioned in the introduction. Namely, I add the *information* and *cultural* distance measures subsequently into the regressions so as to disentangle the impact of the geographical distance on international asset holdings into an information-based and a cultural component.

While the set of host country is initially the same for all source countries, not every investor country holds a positive amount of foreign equity in a given host. In other terms, while a source

country A might have invested in destinations C and D, it might also be the case that another source country B holds foreign portfolio assets only in country C and completely neglect country D. A possible way to deal with such zero observations across different subjects is to specify the dependent variable as $\log(1 + ASSETS_{ij})$. Doing so, whenever there is a missing observation between a given pair of source & host, one obtains zeros which would then be included into regressions. However, there is a large gap between the number of available observations for the dependent variable and the number of observations for different variables and filling up with zeros the cases for which there are no bilateral assets holdings, would significantly alter the distributional properties of the underlying model. To avoid such biases, I drop zero observations from the raw dataset for the corresponding set of exogenous and endogenous variables when working with a linear model specification.[14] I also exclude all bilateral equity holdings involving a destination classified as an "offshore financial center". Although the amounts invested in OFCs represent non-negligible portions of reporting countries' total foreign equity assets, the underlying motivation follows Lane and Milesi-Ferretti's statement (2008) who note that these destinations act *as pure intermediaries, and are neither true sources nor final destinations of investment* (p. 543-544).[15]

The empirical analysis consists of estimating the empirical counterpart of the theoretical relationship given in equation (3.1) and a number of its variants such as

$$\log(ASSETS_{ij}) = c + \alpha \log(SIZE_{ij}) + \beta \log(DISTANCE_{ij}) + \gamma \log(RET_j) + \delta' \boldsymbol{Z}_{ij} + \varepsilon_{ij} \tag{9}$$

where $DISTANCE_{ij}$ is the distance between the source and destination countries, RET_j is the host destination country's market return, $SIZE_{ij}$ is the product of source and host countries' GDPs, and $\boldsymbol{Z}_{ij}$ a set of additional control variables that will be introduced later. To gauge for the impact of

[14] On the other hand, non-linear Tobit regressions allow one to include zero observations for the dependent variable by censoring data on the left.

[15] I exclude from the original data set the following fifty OFC destinations: Andorra, Anguilla, Antigua and Barbuda, Aruba, the Bahamas, Bahrain, Barbados, Belize, Bermuda, British Indian Ocean Territory, Cayman Islands, Cook Islands, Costa Rica, Cyprus, Dominica, Dominican Republic, Falkland Islands, Faroe Islands, Gibraltar, Grenada, Guernsey, Haiti, Hong Kong SAR of China, Ireland, Isle of Man, Jersey, Lebanon, Liechtenstein, Luxembourg, Macao SAR of China, Malta, Marshall Islands, Mauritius, Monaco, Netherlands Antilles, Niue, Panama, Puerto Rico, Reunion, Samoa, Seychelles, Singapore, St. Kitts and Nevis, St. Vincent and Grenadines, Switzerland, Turks and Caicos Islands, United States minor outlying islands, Vanuatu, Virgin Islands (UK) and Virgin Islands (US).

countries' economic masses, I use nominal GDP data from the World Economic Outlook database. Given the panel feature of the data set, the model specification depends on the structure imposed to the residuals. As a first alternative, I start by estimating the model with random effects on both sides as specified in (3.4). Second, to control for unobservable source country-specific factors, I add fixed-effects and estimate the following equation

$$\log(ASSETS_{ij}) = c_i + \alpha \log(SIZE_{ij}) + \beta \log(DISTANCE_{ij}) + \gamma \log(RET_j) + \delta' \mathbf{Z}_{ij} + \varepsilon_{ij} \quad (10)$$

Data for bilateral equity holdings comes from the International Monetary Fund's "Coordinated Portfolio Investment Survey" database available at the IMF's web site. Originally, the CPIS data is a multidimensional array representing the reporting countries' bilateral foreign holdings (organized in columns) across the same set of 237 destinations (organized in rows).[16] From the original data matrix, I consider a subset consisting of the bilateral equity investments reported by 24 developed and emerging-market source countries in various destinations.[17] Nevertheless, the sample is representative enough so that the sum of the foreign assets held by this subset of countries corresponds to 72% of the 74 reporting countries' total foreign assets as provided in the CPIS 2006 data files.

To proxy for the international market frictions, I first employ $Geographical\ Distance_{ij}$, defined as the flight distance in kilometers between the source and host countries' capital or main cities. It is known from previous literature on international goods and asset trade that the geographical distance stands as a good proxy for the bilateral frictions across national markets. Concerning physical goods, the impact of the geographical distance is easy to understand: Distance can proxy for, say, transportation costs across boundaries. However, assets are weightless and trading in assets should not be affected by the physical proximity. Yet, evidence show that this is not the

[16] I also exclude three destinations labeled as "international organizations", "other countries (confidential data)", and "other countries (unallocated)". The data loss due to this elimination, is trivial: The sum of the foreign assets held in these destinations does not exceed 15% of the total assets reported for a particular country in the group of developed countries (the maximum observed for Australia), and 17% in the group of emerging countries (the maximum observed for Poland).

[17] In the group of developed-market source countries we include Australia, Canada, Finland, France, Germany, Italy, Japan, the Netherlands, Spain, Sweden, the UK and the USA; while the group of emerging-market source countries consists of Argentina, Brazil, Czech Republic, Hungary, India, Indonesia, Malaysia, Mexico, Poland, South Africa, South Korea and Turkey.

case. In many empirical studies, the coefficient estimates on the geographical distance is shown to be statistically significant, suggesting that distance plays a key role to explain the international capital flows or assets. That is said, the literature lacks of a rigorous treatment on whether the effect of the distance on asset holdings or trades, is essentially related to some informational or a behavioral aspect of the investor's portfolio selection and decision-making processes. I argue that the effect of the physical distance on the geographical patterns of cross-border asset holdings can be decomposed into two components, i.e. "information distance" and "cultural distance", which account respectively for information- or familiarity-based aspects of investors' preference towards geographically proximate markets. The following sections briefly discuss these two alternative measures of distance and describe the methodologies employed.

3.2 Cultural Distance & Information Distance

3.2.1. Cultural Distance

The first measure proposed to substitute the geographical distance is the cultural distance. Earlier studies have provided valuable insights into the economic outcomes of individuals' or societies' cultural characteristics using various dimensions to control for such effects. For example, religion is an important aspect likely to shape out societies' and corporations' culture. Guiso *et al.* (2003) use the World Values Survey to identify the relationship between intensity of religious beliefs and economic attitudes. They find that religious beliefs are associated to good economic attitudes that are more favorable to higher per capita income and growth. Using data on Finnish investors' shareownership and equity trades, Grinblatt and Keloharju (2001) point out to the importance of language stating that *investors whose native tongue is Finnish prefer to hold and trade in Finnish companies that publish their annual reports in Finnish to Finnish companies that publish their reports in Swedish and vice versa* (p. 1054). Stulz and Williamson (2003) note that *if, as argued by the practitioners of behavioral finance, individuals have psychological biases that matter for finance, it would be surprising that individuals' view of the world as determined by their culture does not matter for how they view and act in financial markets* (p. 347). Thus, culture provides a suitable framework to control for various behavioral

biases advocated by the practitioners of behavioral finance and which would ultimately shape out the observed geography of international portfolio holdings.

To develop a quantitative measure of market imperfections due to culture, I make use of Geert Hofstede's seminal work on cultural affinities. Originally, the study consists of an analysis of a large data base of employee values scores collected by IBM between 1697 and 1973 in more than 70 countries. Since 2001, Hofstede's work lists cultural dimensions scores across 74 countries and regions partly based on replications and extensions of the initial IBM study. Based upon a model that identifies the primary dimensions to describe a country's cultural patterns, Hofstede distinguishes five different indicators. I include four indicators however, due to limited data availability for the fifth one, which is the "long-term orientation". Specifically, I employ: 1) the Power Distance Index (PDI) which represents the extent to which the less powerful members of organizations and institutions accept and expect that power is distributed unequally. This index suggests that a society's level of inequality is endorsed by the followers as much as by the leaders. 2) Individualism (IDV) represents the degree to which individuals are integrated into groups. On the individualist side, we find societies in which the ties between individuals are loose. On the collectivist side, we find societies in which people from birth onwards are integrated into strong, cohesive in-groups. 3) Masculinity (MAS), the opposite being femininity, refers to the distribution of roles between genders. In countries where masculinity is strong, people are keener to seek competitive outcomes; managers tend to make decisions on their own (De Jong and Semenov, 2002). 4) Uncertainty Avoidance index (UAI) which deals with a society's tolerance for uncertainty and ambiguity. It indicates to what extent a culture programs its members to feel either uncomfortable or comfortable in unstructured (novel, unknown, or surprising) situations. Uncertainty avoiding cultures try to minimize the possibility of such situations by strict laws and rules; safety and security measures. The next table shows the original numerical scores associated to these dimensions.

[Insert Table 1 here]

Initially, the raw data set comprises 60 countries plus 3 geographical regions in which several

countries are reported as sharing the same estimated scores. In particular, 7 countries are included in the group "Arab World", 4 countries are included in the group "East Africa", and 3 countries are included in the group "West Africa". The column labeled "C." displays a particular country's cluster membership that I obtain as a result of a four-means cluster analysis over the scores associated to these four dimensions reported in the table. In these terms, we observe that while France, Germany and Italy fall within the same cultural cluster, the US, UK and Australia share the common cultural cluster, which suggest the relative accuracy of Hofstede's cultural dimension scores to distinguish between different societies' cultural affinities. Based upon this multivariate data matrix consisting of 74 countries' individual scores on 4 different subjects, I define the *cultural distance* between two countries as follows,

$$(11) \qquad Cultural\ Distance_{ij} = \sum_{c=1}^{4} \left[\left(S_{ci} - S_{cj} \right)^2 / V_c \right] / 4$$

where S_{ci} is the score for the cth cultural dimension of the country i. Data on individual countries' cultural dimensions comes from Hofstede's work on cultural affinities as emphasized above. V_c is the variance of the cth cultural dimension across all countries in the sample. The formulation above is due to Kogut and Singh (1988) and frequently employed in applied work (see, among others, Beugelsdijk and Frijns, 2010; Lucey and Zhang, 2010; Reus and Lamont, 2009; Lee *et al.*, 2008).

3.2.2. Information Distance

The second measure proposed to substitute the geographical distance is the information distance. In contrast to cultural distance however, existing studies have not come up with such a direct quantitative measure of information distance even if imperfections linked to informational asymmetries are widely recognized as a key determinant of the geographical patterns of cross-border portfolio holdings. In related empirical work, researchers have proposed, beside the physical distance, several other proxies to control for information-based frictions. Coval and

Moskowitz (1999) suggest that air fares or phone rates can be good substitutes to gauge for the economic distance between countries. Ahearne *et al.* (2004) point out to the information content inherent in US investor protection regulations, i.e. accounting standards, disclosure requirements and regulatory environment, as an important factor to explain the home bias. Sarkissian and Schill (2004) note that economic distance can help to explain overseas listing decisions of firms, and that cross-listing is more frequent across markets that are geographically and industrially close to each other. According to Bekaert (1995), inefficient settlement systems and poor accounting standards can be at the root of informational barriers against equity flows into emerging markets. Portes and Rey (2005) capture the informational dimension inherent to cross-border portfolio flows by using bilateral telephone calls, the number of bank branch subsidiaries, and the degree of overlap in trading hours across countries.

To develop a quantitative measure of the information distance between two countries, I employ an indirect theoretical framework developed by Cilibrasi and Vitanyi (2007). Omitting technical details which would overcome the scope of the present paper, the authors develop a theory of semantic distance between a pair of objects and propose the following formula of the *Normalized Information Distance* between two objects x and y,

$$NID_{xy} = \frac{K(x,y) - min\big(K(x), K(y)\big)}{max\big(K(x), K(y)\big)} \tag{12}$$

where $K(x)$ is the *Kolmogorov complexity* of the string x which refers to the *length* of the shortest computer program of the fixed reference computing system that produces x as the output. However, the expression above is uncomputable since Kolmogorov complexity $K(x)$ is uncomputable. Cilibrasi and Vitanyi (2007) apply this theory to construct a formula that extracts a measure of distance between different objects within the World Wide Web, which is undoubtedly the largest database ever created.[18] Namely, they define the following *Normalized Google Distance* between two strings x and y,

[18] The authors' insight can be summarized as follows. Words acquire their meaning from the way they are used in the society and, for computers the equivalent of "society" is "database", and the equivalent of "use" is "way to search the database".

(13) $$NGD_{xy} = \frac{max\{\log f(x), \log f(y)\} - \log f(x,y)}{\log N - min\{\log f(x), \log f(y)\}}$$

where $f(x)$ denotes the number of web pages containing the string x, and $f(x,y)$ denotes the number of web pages containing both x and y simultaneously, as reported from searches performed using Google. Finally, N corresponds to the cardinal of a universal set including all web pages listed within Google. Note further that the results are insensitive to N which can be arbitrarily chosen with the unique condition of being sufficiently larger than $max\{f(\cdot)\}$.

In this paper, I employ the previous formulation of the *Normalized Google Distance* to derive a quantitative measure of the information flow between countries. To obtain the necessary inputs for calculations, I use data on bilateral telephone traffic as a proxy for the overall information flow between two countries. Portes and Rey (2005) also use telephone traffic data as a direct measure of information exchange between countries by normalizing the volume of telephone calls from country i to country j by the square root of the product of their respective GDPs. Unlike Portes and Rey (2005) however, I implement a different approach to control for the information distance that is analogous to those proposed by Cilibrasi and Vitanyi (2007). Namely, I define the information distance between two countries i and j as follows:

(14) $$Information\ Distance_{ij} = \frac{max\left\{\log\left(\frac{f(s)}{p_s}\right), \log\left(\frac{f(d)}{p_d}\right)\right\} - \log\left(\frac{f(s,d)}{p_s + p_s}\right)}{\log N - min\left\{\log\left(\frac{f(s)}{p_s}\right), \log\left(\frac{f(d)}{p_d}\right)\right\}}$$

where $f(s)$ and $f(d)$ denotes country s and d's total phone traffic (incoming and outgoing calls) over a given time period, $f(s,d)$ denotes the bilateral phone traffic involving countries s and d, all these terms normalized by the countries' respective populations, p_s and p_d. Finally, I substitute N by the total telephone traffic over the full data set which yields to a number sufficiently larger than any conceivable $f(\cdot)$. Implementing the previous formulation instead of directly using the volume of the bilateral telephone traffic between two countries (either in levels or in logs) can be justified given the objective to obtain a *distance* measure. As underlined by Cilibrasi and Vitanyi (2007), equations (13) and (14) have several interesting numerical properties. For example, under

this formulation, the information distance between two distinct objects is bounded on the continuous interval $[0, \infty)$ and is always nonnegative. By definition, one obtains $Information\ Distance_{ij} = 0$ if $i = j$, and $Information\ Distance_{ij} = Information\ Distance_{ji}$, i.e. the distance between two objects is symmetric. The measure is also scale invariant in the sense that the set of different outcomes does not depend on the number N. In fact, as N grows, the relative frequencies of different $f(\cdot)$ tend toward a fixed fraction of N. For calculations, I obtain data on international telephone traffic in millions of minutes over the year 2006 and countries' populations in millions as of the end-2006 from Telegeography. The full data set comprises a total of 1545 observations regarding the bilateral telephone communications between a set of 102 source and 206 destination countries. The total international telephone traffic during 2006 is 378 474 millions of minutes of which about 66 568 millions originate from the traffic between Canada, Mexico and the United States. To avoid measurement biases, I normalize the volume of the international telephone traffic by the countries' population. In fact, examining 2006 data, we observe that the total telephone traffic between France and Italy and is about 1340 millions of minutes while the volume of the telephone traffic between France and Belgium is slightly lower, about 1 310 millions of minutes. However, looking at these numbers we can't deduce that the communication between Italy and France is more intense than between Belgium and France since Italy's population is about 6 times than that of Belgium's.

3.2.3. Information and Cultural Distance vs. Geographical Distance

Given the respective definitions of information and cultural distance measures, how one can assess their relationship with the geographical distance? In order to justify the central hypothesis that the effect of the geography on bilateral portfolio holdings can undergo an informational and/or a cultural component, these two measures should not be strongly correlated with the geographical distance. As a simple response to such concerns, the figure below provides the scatter plots of cultural and information distance measures against the geographical distance.

[Insert Figure 1 here]

The graph in the upper panel shows the scatter plot between geographical distance (plotted on the horizontal axis) and information distance (plotted on the vertical axis). The graph in the lower panel shows the scatter plot between geographical distance (plotted on the horizontal axis) and cultural distance (plotted on the vertical axis). For purposes of comparability, I use standardized scores with zero mean and unit variance for each variable. Looking at Figure 1, we observe that these two artificial distance variables can effectively capture distinct aspects of the international capital market frictions by disaggregating the unified effect of the physical distance. The rank correlation between cultural distance and geographical distance is about 8% while the rank correlation between information distance and geographical distance is a bit higher but still weak, by about 32%. Indeed, countries that are informationally or culturally close (resp. remote) to each other are not necessarily those that are geographically close (resp. remote). For instance, while United States and Canada rank 59th regarding the physical distance between their main cities, these two countries rank 33th and 27th respectively on the basis of information and cultural distance measures respectively. Another striking example with respect to the irregularity between geography and cultural affinities involves the US and Australia who rank 2nd according to cultural distance measure although these two countries take the 1902th place in the sample once we consider the flight distance about 16 000 kilometers between Sydney and New York.

3.3 Additional Controls

Beside the core variables identified above I also conjecture that the foreign equity holdings would depend on a number of additional gravity-type variables as proposed by previous studies. Specifically, I hypothesize that foreign openness, economic development, and market transparency can influence investors' decision to allocate their foreign assets into a particular host market. To control for these effects, I estimate augmented forms of the initial gravity model by adding an array of additional variables described below.

3.3.1. Economic Development

I test whether the level of the destination country's economic development is considered as an asset for investors. As noted by Chan *et al.* (2005), a country's degree of development and market sophistication should presumably have a positive impact in attracting foreign capital.[19] At an individual investor scale, Grinblatt and Keloharju (2001), and Karlsson and Norden (2007) argue that home bias is inversely linked to investors' sophistication level. Additionally, as emphasized by La Porta *et al.* (1997, 2000), investors may also feel more comfortable when they purchase stocks from developed markets with healthier legal systems and better investor protection rights. In line with these studies, the hypothesis maintained is that the level of economic development and market sophistication could lower the costs due to market frictions via better structured and developed financial markets.

I dispose several proxies to capture the impact of the economic development on bilateral equity positions. First, I propose to introduce the destination country's GDP per capita, $GDPPC_j$. The second variable measures the degree of the recipient country's financial market sophistication. Namely, the variable $FMSOP_j$ is an index variable with scores given out of an upper limit by 7. Higher index values are assigned to superior market sophistication for the country in question. The third variable, i.e. the investor protection index ($INVPRO_j$) is an aggregate measure combining a country's 1) degree of transparency in transactions; 2) the liability of self-dealing, and 3) the shareholders' ability to sue officers and directors for misconduct.[20] For countries with better investor protection standards, the index assigns higher values. Given their definitions, all these controls are expected to be positively correlated with the size of bilateral portfolios.

3.3.2. Foreign Openness

Examining the impact of countries' cultural backgrounds on cross-sectional differences in investor protection standards, Stulz and Williamson (2003) find that *finance is more valuable to countries that can benefit more from being open to international trade* (p. 338). According to

[19] While Chan *et al.* (2005) consider the economic development, the stock market development and the investor protection standards as separate categories of explanations, I believe that they can all be embedded into a single group, capturing various aspects of a particular country's degree of economic development.

[20] In line with the *protecting investors* dimension of the World Bank's *Doing Business Project*, Djankov *et al.* (2008) propose a measure of legal protection of minority shareholders against self-dealing transactions benefiting controlling shareholders.

Ahearne *et al.* (2004), although capital controls have been reduced, they can still affect cross-border investment and the authors give the example of US investors who underweight the Chinese market maintaining substantial barriers to foreign investment. As such, the openness of a country's international trade or capital flows may promote the foreign investment inward, whereby affecting the geographical spread of the source country's target foreign destinations. The degree of openness can also be considered as a proxy for the information cost that investors bear whenever they consider investing into a particular foreign market. To control for these effects, I employ two variables. The first one is an artificial variable assessing a country's degree of openness to international trade. Namely, $FTRADE_j$ corresponds to the volume of foreign trade in goods and services (i.e. the nominal sum of the imports and exports) scaled by the country's nominal GDP for the year-end 2006. The second variable to control for the effect of foreign openness is mergers & acquisitions, MA_j, is a measure of the foreign openness.[21] It is the monetary sum of a country's cross-border mergers & acquisitions (i.e. purchases plus sales) scaled by the GDP. Again, I expect both of these controls to enter the model with a positive sign.

3.3.3. Familiarity

Ricciardi (2008a) defines "familiarity" as an *inclination that alters an individuals' perception of risks implying that investors tend to fear less from familiar risks than those that are unfamiliar* (p. 101). With respect to the international portfolio allocation and the home bias literature, the concept of familiarity implies that investors shrink their portfolios across investment alternatives they feel more familiar with. To gauge for this impact, previous studies employed so far different proxies such as geographical proximity, trade linkages or cultural affinities. For example, Massa and Simonov (2006) argue that *investors prefer to invest in countries geographically and professionally closer to their domestic country* (p. 634). Coval and Moskowitz (1999) note that US fund managers exhibit systematic biases toward nearby firms' stocks, and Huberman (2001) provide similar evidence on the effect of geographical proximity using a sample US households' stock holdings. By the same token, other studies proposed that cultural affinities such as the

[21] Yet, this study is also the first to introduce this variable to gauge for a country's foreign openness.

existence of a common language (Grinblatt and Keloharju, 2001)[22], the origin of the legal system (La Porta *et al.*, 1997; La Porta *et al.*, 2000), the intensity of bilateral trade (Chan *et al.*, 2005), the religious connections (Stulz and Williamson, 2003)[23] etc. would also influence portfolio selection.

Following these studies, I use an array of explanatory variables to control for possible familiarity effects in bilateral equity holdings. The first measure, $BTRADE_{ij}$, is the relative amount of country i's foreign trade due to transactions with country j as the partner. More explicitly, a percentage of, say, 10% between a source country A and a destination country B, means that the bilateral imports and exports transacted with country B corresponds to one tenth of the sum of all imports and exports of the country A. I also add a number of indicator variables relative to other aspects of familiarity. Namely, $LANG_{ij}$, is a dummy variable equal to 1 if the source and host countries share a common language family and 0 otherwise. The second variable, $LEGAL_{ij}$, is a dummy equal to 1 if the source and host countries' legal systems derive from a common origin.[24] The third variable REG_{ij}, is a dummy variable equals to 1 if the source and host countries fall into the same geographical region. I compute this binary variable by distinguishing between four broad geographical regions, i.e. Africa, America, Asia-Pacific and Europe following the classification proposed by the World Bank.

3.3.4. Transparency

This category of controls is motivated from recent evidence set forth by Gelos and Wei (2005) who investigate the effect of a country's transparency on the foreign investment patterns of emerging market funds. The authors find that "international funds prefer to hold more assets in more transparent markets" and "both government and corporate transparency have separate and distinct positive effects on investment flows from international funds into a particular country". Compiling various sources, they elaborate two groups of indicators, namely "government opacity"

[22] Grinblatt and Keloharju (2001) argue that Finnish households are more likely to invest in the stocks of Swedish firms communicating in the investor's native tongue.

[23] Stulz and Williamson (2003) identify three channels through which culture can affect finance: *First, the values that are predominant in a country depend on its culture. […] Second, culture affects institutions. For instance, the legal system is influenced by cultural values. Third, culture affects how resources are allocated in an economy* (p. 316-317).

[24] I distinguish between legal systems from English, French, German and Scandinavian origins.

and "corporate opacity". In this study, however, I use other proxies directly observable and allowing for larger sample coverage.[25] Implementing Gelos and Wei's analysis to country-level equity positions, I conjecture that more transparency would imply less (perceived) risk and/or weaker information asymmetry and encourage investors to invest more in this market.

The first variable used to assess a country's transparency is the "Corruption Perceptions Index", abbreviated as CPI_j. Briefly, a country's CPI score relates to perceptions of the degree of corruption among public officials and politicians as seen by business people and country analysts. Original country scores range from 0 to 10, with higher values assigned to weaker perceived corruption. Second, I include the "Judiciary Independence Index", denoted $JUDIN_j$, which measures the degree of independence of a country's judiciary system. Raw scores are given in numerical values out of 7. Values closer to 7 are assigned to higher perceived dependence of the judiciary system to political influences of members of government, citizens and firms. The third variable, the "Capital Market Controls" index (CMC_j) directly measures the intensity of capital market controls within a country. It indicates the percentage of capital controls not levied as a share of the 13 different types of international capital controls reported by the IMF. Originally, individual country ratings range from 0 to 10, with lower scores assigned to countries with higher restrictions on foreign capital flows. Therefore, while I expect a positive coefficient estimates on CPI_j, while the variables $JUDIN_j$ and CMC_j should be inversely related to the dependent.

3.3.5. Portfolio Diversification

After all, the basic premise behind diversification is that it ought to enhance either the expected return given the risk or to reduce the risk given the expected return of a portfolio. Thus, the last category of controls checks whether the existing geography of international equity holdings is guided or not by any diversification motive. Doing so, I employ three variables using data from Morgan Stanley Capital International. Using country gross index series (dividends included) in US$ terms over the period 2002:01-2006:12, I construct 1) RET_j, the average 5-year

[25] The corporate opacity measure used by Gelos and Wei (2005) covers 53 countries. The scope of my analysis, however, requires a larger data collection in that we deal with up to 102 destination countries.

return on the country j's stock market; 2) $SHRP_j$, the risk-adjusted-return on country's broad market index computed as the ratio of mean excess return to the standard deviation; and 3) COR_{ij} is the correlation between the country i and j's stock market returns.[26] In light of our theoretical detour above, I employ the average return directly in the gravity model, and I control for the results using the Sharpe ratio and the market correlations later.

3.4 Descriptive Statistics

Table 2 provides summary statistics regarding the variables described above.

[Insert Table 2 here]

For each variable, Table 2 shows the total number of observations, the mean, the standard deviation, the minimum and the maximum scores. The first column displays the expected signs of the regression coefficients. The large differences in the number of available observations across variables imply that regressions are run over an unbalanced panel. Following Papaioannou (2009), I transform raw scores of financial market sophistication, investor protection, foreign trade to GDP, corruption perceptions, judiciary independence and capital market controls so as to restrict scores between 0 and 100. The reason for which the maximum score associated to the financial market sophistication is that descriptive statistics are computed over a filtered data set excluding observations involving an OFC as the destination. The largest bilateral portfolio asset position concerns US investors' holdings in the UK. Germany and Switzerland are the countries closest to each other as shown by the cultural distance. The smallest informational distance is between Japan and South Korea. The smallest and the largest geographical distance are observed between Finland & Estonia and Spain & New Zealand. In addition, while the physical distance between the US and the UK is up to 5500 kilometers long, these two countries are found to be very close to each other according to both the informational and cultural distance measures. Looking at different variables groups, we observe a significant cross-sectional variation. For example, the GDP per capita ranges between a minimum of US$ 122 for Zimbabwe to a

[26] In case there is lack of a directly observed time series for a particular destination, I substitute it by an appropriate regional index.

maximum value of US$ 72,768 for Norway. The UK ranks first in the market sophistication category while New Zealand gets the highest score on the investor protection index. Libya and Swaziland rank the last places with respect the financial market sophistication and the investor protection indices respectively. An interesting observation is related to the bilateral trade between Canada and the US as shown by a maximum score by 69%. This suggests that 69% of Canada's foreign trade volume originates from transactions with the US as the partner. The second highest score concerning bilateral trade is observed between the US and Mexico with a bilateral foreign trade volume about 68%.

For modeling purposes, it is also useful to check for the pairwise correlations between these variables. Overall, correlations between different variables belonging to the same category are not strong enough to raise concerns about multicollinearity. For example, the average correlation between the economic development variables, i.e. GDP per capita, financial market sophistication and investor protection is about 44%; and the average correlation between the variables in the transparency category, i.e. corruption perceptions, judiciary independence and capital market controls is about 49%. Further, the pairwise correlations between the three distance measures are also sufficiently weak: The average is 22%. This suggests that each of these proxies may truly serve to account for diverse aspects of the international market frictions within a gravity model framework.

4 Empirical Analysis

4.1 The Gravity Model of Bilateral Equity Holdings

Tables 3 to 5 present the estimation results for the gravity model as described in equations (9) and (10). The dependent variable is $\log(Assets_{ij})$, i.e. logarithm of the bilateral equity assets held by a source country i in the host country j as of the end-2006. Results for the full sample includes 24 source countries and are reported in panels A of Tables 3 to 5. Besides, I also distinguish two subsamples which include 12 developed-market source countries and 12 emerging-market source countries separately, whose results are reported in panels B and C respectively.

[Insert Tables 3 to 5 here]

In each table, columns (1) to (3) report regression results with geographical, cultural and information distance variables separately to highlight their respective effects on bilateral asset holdings. Alternatively in columns (4) to (6), I control for possible spillovers from information and cultural distances to geographical distance using two-stage regressions to check whether geography affects bilateral asset holdings through information or culture. Concerning Table 5 with non-linear regressions, I make the same comparisons by testing the joint effects of the information and cultural distance on the dependent along the geographical distance, since a two-stage procedure is not available for panel Tobit regressions. To guarantee comparability between the coefficients on different distance variables, I standardize raw scores to bring them to the same scale.[27] Due to the log-log specification imposed to both sides of the regressions, the estimated coefficients are all in terms of elasticities. Because of the large gap in the number of available observations across different variables in the dataset, in particular concerning information and cultural distance measures (475 against 1560), I eliminate the cases where there is no observation for information distance.[28]

I consider various estimation techniques. In Table 3, regressions are estimated using generalized least squares (GLS) in columns (1) to (3), and generalized two-stage least-squares (G2SLS) in columns (4) to (6). GLS estimations assume a two-way random effect specification for residuals. I check for the adequacy of random effects for GLS estimations via Breusch-Pagan Lagrange multiplier test (see Breusch and Pagan, 1980) where the null hypothesis is that variances across entities are zero. Breusch-Pagan LM statistic is asymptotically distributed as a chi-squared distribution and the resulting χ^2 values are all significant at 1% level suggesting that a random effects specification for error terms could be appropriate. I also verified the results using emerging and developed countries subsamples. The results remain unchanged. Regarding the G2SLS procedures, first-stage regressions include the geographical distance as the dependent

[27] Geographical distance is in kilometers whereas information and cultural distance variables are without unit.

[28] Interestingly, after performing this elimination, I am left with fewer observations for the cultural distance than for information distance. Overall, the data set still comprises a sufficient number of observations to perform necessary calculations.

and information and cultural distances as predictors in columns (4) to (6) where I report the coefficient estimates of information and cultural distance variables obtained from the first-stage regressions. Table 4 assumes a one-way fixed-effect model to estimate the gravity model whereby I include dummies to control for unobservable source-country specific effects. As in the case for random effects under the GLS estimations, I also control for the adequacy of fixed-effects using a joint test to see if source country dummies are all equal to zero. As suggested by highly significant F statistics, I find that adding fixed-effects across different entities can also be appropriate to control for unobservable country-specific characteristics. In Table 4, columns (1) to (3) display the results obtained from panel OLS procedures where geographical, information and cultural distance variables enter the regressions separately. As in Table 3, columns (4) to (6) allows information and cultural distance measures to influence geographical distance indirectly through instrumental variables estimations.

In Table 5, I opt for a non-linear procedure and estimate the gravity model using panel Tobit regressions. In fact, previous estimations set the dependent variable as log of the bilateral equity assets, thereby discarding several observations with zero cross-border assets between a given source and host country pair i,j. Alternatively in Table 5, following Lane and Milesi-Ferretti (2008), I set the dependent as $\log(Assets_{ij} + \varepsilon)$ with ε being a small number and use the resulting zero observations within a panel Tobit specification with left-censored data. For comparison, while regressions reported in columns (1) of Table 3 and 4 comprise 383 observations, Tobit regression with the same set of predictor variables employ 421 observations. For each model specification the same set of destinations is used depending on data availability. In the end, even though the underlying econometric theory and model assumptions differ across estimations, both statistical and economic implications of the estimated coefficients point out to interesting results, regrouped in three categories.

First, consistent with previous studies such as Papaioannou (2009), Lane and Milesi-Ferretti (2008) or Portes *et al.* (2001), we observe that a gravity model of international trade in financial assets performs fairly well. Regressions capture a more or less important part of the variation in

data as suggested by the goodness of fit measures: Concerning the panel OLS regressions with source-country fixed effects, the estimated model including return, size and geographical distance as predictors explains 64% of the cross-sectional variation in asset holdings over the full sample. A non-linear specification using left-censored data with the same set of predictors also captures almost half of the variation in the data (the pseudo R-squared close to 50%). In contrast to results presented in panels C where I include emerging countries into the estimations, the model fit is improved when I restrict the estimations within the developed countries subsample; the adjusted R-squared is above 70% through columns (1) to (3) in Table 4.[29] Looking at regression coefficients, we observe that size and distance variables are appropriately signed through different samples and specifications. Bilateral asset holdings respond positively to source and host countries' economic mass as suggested by highly significant coefficients. Geography has a counter-intuitive impact on bilateral asset holdings, which tend to decrease systematically as the physical distance between the source and destination countries increases. Further, physical distance affects the behavior of investors in emerging countries more than it does for investors in developed countries as suggested by the large gap between the estimated coefficients on geographical distance reported in panels B and C of the tables: Under the panel OLS specification we observe that when geographical distance is cut by 50%, emerging countries' foreign equity assets is more than doubled. This observation also underlines the importance of the gravity model to understand the patterns of international portfolio investments. From a purely diversification perspective, remote geographies should be particularly preferred by investors willing to diversify away the risk on their portfolios, implying that foreign holdings should increase with distance. However, investors do merely follow this rule and tend to concentrate their investments across a limited set of foreign countries closer to their own home country.

Second, beside the well-documented impact of physical distance on international equity investments, the information and cultural distance measures also perform quite well in a gravity equation. This suggests that these two alternative distance measures are also able to account for

[29] A similar result is found by Lane and Milesi-Ferretti (2008) who also distinguish between regression results using developed countries and emerging countries separately.

other aspects of international market frictions beside the geographical distance. Using the full sample and developed-market source countries subsample, both information and cultural distance enter the regressions with well-determined coefficients significant at the 1% level. Using the emerging-market source countries, however, both distance measures are no longer statistically significant and the cultural distance measure is positively signed implying that it is economically insignificant too. Looking at the estimated coefficients on distance proxies however, we can deduce that the effect of the geographical distance on bilateral assets is systematically higher than those of the information and cultural distance since regressions are run using standardized scores. For example, under the GLS specification, we observe that while the geographical distance has a coefficient by –0.88 in panel A, the coefficients on information and cultural distance variables are –0.50 and –0.32 respectively. Given the log-log specification, this implies that a 10% decrease in information and cultural distance between two countries is expected to increase bilateral equity holdings by 5.4% and 3.5% respectively.

Third, empirical evidence suggest that the controversial effect of the geography on international portfolio investments is more likely to undergo an information-based channel rather than reflecting investors' preference towards foreign markets they feel more familiar with and culturally closer to their own home market. In fact, inspection of the results obtained from two-stage procedures given in columns (4) to (6) allows one a direct assessment of the separate and joint impact of the cultural or information distance first on geographical distance and second on bilateral equity assets. Looking at panel A of Table 3, we observe that while the first-stage coefficients associated to both information and cultural distance measures are separately significant in columns (4) and (5), we note that cultural distance is no more significant once it is used to predict the geographical distance jointly with the information distance. The same result also holds for the first stage estimation results under the fixed-effects specification in Table 4: Information distance explains geography more than the cultural distance as suggested by the significant coefficient for the first variable and insignificant for the second. Further, comparing the respective coefficient estimates of the first-stage regressions reported in columns (4) and (5), we also note systematically higher coefficients associated to information distance than the one

obtained for cultural distance, an observation that also holds for Tobit regressions displayed in Table 5. Specifically, we note that the impact of the geographical distance on the dependent variable is considerably reduced when information distance is included into the regressions. Thus, evidence presented so far tend to support Massa and Simonov's (2006) view that *familiarity-driven investment is a rational response to information constraints as opposed to a behavioral heuristic* (p. 634), and thereby the home bias puzzle and investors' preference towards physically markets are part of an information-based story as opposed to psychological biases argued by the practitioners of the behavioral finance.

4.2 Regressions with Other Variables

In this subsection, I augment the baseline specification by adding into the initial model a number of additional variables previously employed by the related literature. Namely, I introduce several proxies to control for the likely effects of economic development, familiarity, foreign openness and transparency on sample countries' foreign equity holdings. I also check for any diversification motive in the observed geography of cross-border equity holdings via two proxies, namely, the destination market's Sharpe ratio and the bilateral correlation coefficient between the source and host markets. For modeling purposes, omitting the return component and with a bit of rewriting, I consider the following specification for the baseline gravity model as proposed by Portes and Rey (2005, p. 275),

$$\log(Y_{ij}) = c + \alpha \log(M_i M_j) + \beta \log(\theta_{ij}) \quad (15)$$

Decomposing the log-product yields to,

$$\log(Y_{ij}) = c + \alpha_1 \log(M_i) + \alpha_2 \log(M_j) + \beta \log(\theta_{ij}) \quad (16)$$

As underlined by Portes and Rey (2005) theory suggests that $\alpha_1 + \alpha_2 = 1$ (p. 276). Denoting by $\boldsymbol{Z}_{ij}$ the set of other controls to be included, equation (15) can be rewritten as follows,

(17) $$\log\left(\frac{Y_{ij}}{M_i M_j}\right) = c + \beta \log(\theta_{ij}) + \delta' \mathbf{Z}_{ij} + \varepsilon_{ij}$$

Substituting the terms Y_{ij}, $M_i M_j$, θ_{ij} and $\mathbf{Z}_{ij}$ by their data counterparts, the general estimating model is defined as follows,

(18) $$\log\left(\frac{Assets_{ij}}{GDP_i \times GDP_j}\right) = c_i + \beta \log(Distance_{ij}) + \delta_1(Economic\ Development) + \delta_2(Familiarity) + \delta_3(Foreign\ Openness) + \delta_4(Transparency) + \delta_5(Portfolio\ Diversification) + \varepsilon_{ij}.$$

To estimate (18), I adopt a one-way fixed-effects specification to control for unobservable source country-specific characteristics as this is the case in Table 4. As emphasized previously, either Breusch-Pagan LM test for random effects or the joint *F*-test for source country fixed-effects cannot reject the hypothesis that the model specifications proposed in Table 3 and Table 4 (columns 1 to 3) is appropriate. In fact, this choice is rather motivated by the fact that most of the available controls reflect destination country-specific characteristics such as the financial market sophistication or corruption perceptions, except the familiarity variables such as the bilateral trade or common legal origin dummy. Tables 6 to 8 present the results of estimating the preceding equation by substituting the $DISTANCE_{ij}$ by the geographical, information and cultural distance respectively. In column (1) through Tables 6 to 8, I present the results using the full set of controls. In columns (2) to (6), I present the results obtained by regressing each group of controls separately on the dependent variable, which is specified henceforth as log of the normalized bilateral equity assets, i.e. $\log(Assets_{ij}/GDP_i \times GDP_j)$.

[Insert Tables 6 to 8 here]

In line with previous results provided by Chan *et al.* (2005) and Portes *et al.* (2001), the recipient country's economic development in attracting foreign capital investments inward is well captured, especially via destination countries' GDP per capita variable which is significant and correctly signed whether it is regressed jointly with geographical, information or cultural distance in columns (2). However, unlike Portes and Rey (2001) who report a significant positive impact of

the market sophistication on cross-border equity flows, the two other proxies for countries' economic development, i.e. financial market sophistication index and investor protection index, perform rather poorly with statistically insignificant coefficients. As a further check, I also controlled for the effect of market sophistication and investor protection on bilateral equity holdings separately along with the geographical, information and cultural distance measures in other regressions. I find that both variables enter the models with expected signs and significant coefficients except for the case where I use investor protection index and cultural distance as predictor variables.

Consistent with previous research, familiarity also helps explaining bilateral investments. The respective effects of familiarity proxies used in Tables 6 to 8 are mainly captured by the bilateral trade variable since dummies for common language, legal origin and geographical region are of no economic consequence as shown by statistically significant but negative coefficient estimates regarding the legal origin and geographical region dummies. Indeed, there is no reason to expect bilateral equity assets to decrease when the source and destination countries share a common legal system origin. The statistical insignificance of these binary variables is also quite surprising since prior studies find strong links between these variables and the foreign investment patterns. For example, Aviat and Coeurdacier (2007) report significant positive effect of sharing a common legal origin system on bilateral banking claims. Lane and Milesi-Ferretti (2008) note that speaking a common language is expected to raise bilateral equity holdings by about 50%. A possible interpretation for the failure of these dummies would be that their impact can be soaked up by the bilateral trade which enters the models with very well-determined coefficients. To check for this effect, I have also run the regressions in columns (3) without including the bilateral trade. For example, using geographical distance and dummies for common language, legal origin and geographical region within the same model, I find that only language (beside the geographical distance) is correctly signed with a coefficient significant at 5% level. The same observation also holds for the common language dummy when regressed together with the information distance.

Destination countries' degree of openness to international trade affects bilateral equity

holdings mainly through the mergers & acquisitions variable MA_j which measures the monetary sum of total mergers and acquisitions over the year 2006 scaled by the source country's GDP. Regardless of the distance measure employed in estimations, the coefficient on mergers & acquisitions is statistically significant at the 1% level and correctly signed. I have also used this variable without including foreign trade into different models with geographical, information and cultural distance and I still have found significant coefficients at the 1% level. On the other hand, the countries' foreign trade volume to GDP ratio also enters regressions with statistically significant coefficient estimates, but not with the expected sign: The magnitude of bilateral equity assets between a given pair of source and host countries is expected to respond positively to the host country's level of openness to international trade as proposed by previous studies like Stulz and Williamson (2003). Indeed, the extent to which a given market or country ties economic relations abroad is likely to improve the amount and the quality of information exchange among the partners involved, and consequently, to boost bilateral portfolio investments.

The positive effect of a host country's transparency is on bilateral equity assets well captured in particular by the corruptions perceptions index since the remaining two proxies fail to adequately enter the regressions. Across Tables 6 to 8, the coefficient on the CPI_j is around 0.30 suggesting that a 50% improvement in a given country's degree of corruption could boost bilateral foreign equity inflows by about 15%. Thus evidence broadly support Gelos and Wei's (2005) view that improving a country's transparency could lead to an increase in investment flows, implying a weaker home bias by foreign investors against the country of interest.[30] On the contrary, the two other variables, i.e. the judiciary independence and capital market controls indices, are of no consequence on the dependent variable as shown by insignificant coefficient estimates. As before, I checked for the separate effect of these proxies in other regressions, the results are unaffected either for judiciary independence or capital market controls.

Finally, in columns (6) I control for whether the observed bilateral equity holdings are driven by investors' diversification motive and add two variables into the baseline gravity model

[30] Gelos and Wei (2005) suggest that *becoming more transparent can be an effective way for countries to benefit from international financial integration while avoiding excessive volatility during turbulent times* (p. 3012).

specification. In fact, from a portfolio diversification viewpoint, a destination market's Sharpe ratio should be positively correlated with foreign investors' portfolio assets within this market of interest; and negatively correlated with the pairwise correlation between the source and host countries' market returns since the greater the degree with which two markets are synchronized the lower the benefit from portfolio diversification. However, the results are exactly the opposite of what is expected and further complicate the puzzling picture of international portfolio investments: Not only investors hold fewer assets in destinations with higher risk-adjusted returns as suggested by significant coefficients on the Sharpe ratio; they are also likely to invest more in destinations whose stock markets are more correlated with their domestic stock market: The coefficient on the pairwise correlation between the source and host markets is around 0.63, significant at the 1% level. Previously, Aviat and Coeurdacier (2007) referred to this observation as the *correlation puzzle*. This suggests that risk diversification is an unlikely determinant of bilateral portfolio investments as emphasized by Portes and Rey (2005) who note that investors tend to prefer informationally close markets when they consider investing abroad.

When all predictors enter the regressions simultaneously, only a few of them preserves its explanatory power. First, none of the distance proxies is helpful to explain bilateral equity holdings and their coefficients are greatly reduced according to results displayed in columns (1) of Tables 6 to 8. This suggests that their respective effects are absorbed by other predictors in the model. When I use geographical and cultural distance variables, I observe that bilateral trade, mergers & acquisitions, capital market controls and Sharpe ratio are both statistically and economically significant. The only exception concerns the capital market controls variable whose coefficient loses its explanatory power when regressed along with the information distance. Second, the negative coefficient on the Sharpe ratio generally confirms the estimation results presented in columns (6) where I control if bilateral equity holdings could be driven by a diversification motive. Besides, the coefficients on pairwise market correlation are also positive but not significant at conventional confidence levels. As a further check of the correlation puzzle, I have also run the same regressions with the complete set of predictors by excluding the Sharpe ratio variable. In this case, I have found that COR_{ij} still enters the models with positive estimates

significant at the 10% level. Third, estimations provided in columns (1) further highlight the effect of destination countries' foreign openness on bilateral equity investments from the source markets in the sample as captured by the mergers & acquisitions to GDP ratio, MA_j, which enters the models in all cases whether one uses geographical, information or cultural distance.

Fourth, we observe that bilateral trade is also a major determinant of bilateral equity holdings. Not only it has a systematic positive influence on the dependent variable when it is regressed jointly with other familiarity dummies, but it remains the major predictor in the regressions even though I employ other proxies to control for economic development, transparency, foreign openness etc. To check for the consistency of the effect of bilateral trade on bilateral equity assets, I have run additional estimations for the general model given in equation (18) using different combinations of control variables. Given the unmanageable amount of possible combinations across this set of available variables, I estimated stepwise regressions (both forward and backward) allowing variables to enter the regressions with a significance level equal or more than 10%. Regardless of the distance proxy used in regressions, I find that bilateral trade is systematically the first variable to be included in regressions with well-determined and correctly signed coefficients and that the distance proxies perform rather poorly once bilateral trade makes part of the predictors within the estimated model.[31] Given this systematic effect of bilateral trade on bilateral equity holdings, the relevant question is to ask whether bilateral trade can be considered as a substitute for information or cultural proximity between countries.

4.3 Disentangling the Effect of Bilateral Trade: Information or Culture?

According to Lane and Milesi-Ferretti (2008), there exists a strong link between trade and equity holdings and they suggest that this observation is particularly consistent with the informational potential of the strength of bilateral trade relations across countries (p. 546). They report that a simple econometric specification using bilateral trade as predictor of bilateral equity assets is enough to explain as much as 86% of the variance observed in a data set that covers a subsample

[31] Concerning stepwise procedures, I have used a linear specification for the panel data with source country fixed-effects and the bilateral trade is the first control to enter the models after different dummies introduced for entities.

of OECD countries' foreign equity assets. The strong positive link between bilateral trade and equity holdings would reflect the preference that investors exhibit for "professionally closer countries" as conjectured by Massa and Simonov (2006). Examining the complementarity between bilateral trade and bilateral asset holdings within a simultaneous equation framework, Aviat and Coeurdacier (2007) argue that 1) distance which substitutes transport costs in goods market would induce home bias in asset portfolios and 2) part of the effect of geographical distance on bilateral asset holdings could be related to bilateral trade relations between countries. The authors report that a 10% increase in bilateral trade leads to a 3% increase in bilateral banking claims. Moreover, although the direction of causality between trade and assets runs significantly in both ways, they find that the impact of trade on asset holdings is substantial. Portes and Rey (2005) suggest that information flows via trade in goods could enhance bilateral portfolio holdings to the extent that trading partners are more likely to share and exchange information, thereby eroding a good part of the informational asymmetries in the financial markets.

In light of these studies, I propose a simple way to control for whether the effect of bilateral trade on portfolio holdings goes through our information or cultural distance measures since these two alternative variables are found to be good proxies for geographical distance. Indeed, the direction of causality between trade and asset holdings runs significantly in both ways and omitting bilateral trade in goods within a gravity model of bilateral equity holdings could lead to estimation bias problems as underlined by Aviat and Coeurdacier (2007). Then, it would be conceivable to introduce the variable $Assets_{ij}$ into the basic gravity model given in equation (15) as a plausible predictor of bilateral trade in goods,

$$\log\left(Bilateral\ trade_{ij}\right) = c + \beta \log\left(Distance_{ij}\right) + \gamma \log\left(Assets_{ij}\right) \tag{19}$$

If bilateral trade in goods is endogenous to the baseline gravity model of asset holdings and if the effect of geographical distance on portfolio assets can effectively be decomposed into two components, i.e. information distance and cultural distance, one would expect the predicted values of bilateral trade in goods to be correlated with these two alternative distance measures.

Then, by comparing the correlations coefficients between the predicted values of the dependent variable and different distance proxies, one would understand whether the effect of goods trade on the observed geography of portfolio holdings goes through an informational or cultural channel. I estimate the preceding equation using panel OLS with source country fixed-effects and substitute the terms $Distance_{ij}$ by geographical distance between countries i and j, and $Assets_{ij}$ by bilateral equity assets held by source country i in the host country j scaled by source and host countries' GDPs products. The estimation results are as follows:

$$\log\left(Bilateral\ trade_{ij}\right) = 1.87 - 0.32\log\left(Geographical\ distance_{ij}\right) + 0.50\log\left(Assets_{ij}\right) \quad (20)$$

All coefficients are statistically significant at the 1% level (assuming heteroskedasticity-consistent estimator for the variance) with a satisfactory goodness-of-fit measure of the regression: R^2 of within estimates is close to 0.31. Further, the F test to check if all fixed-effects are jointly equal to zero can be easily rejected as well. Given these results, the next figure provides the scatter plots of the predicted scores for the bilateral trade in goods against the information distance in panel A and the cultural distance in panel B.

[Insert Figure 2 here]

As shown by the linear fits, the scatters broadly highlight the relative importance of information proximity against the cultural proximity to explain the effect of bilateral trade in goods on the geographical patterns of international portfolio investments. Computing the correlation between bilateral trade and information distance, I find a coefficient close to minus 0.38, which is significant at the 1% level. In contrast, the correlation between cultural distance and trade is fairly weak by minus 0.05 and statistically insignificant. In brief, evidence is mostly consistent and provides empirical support to previous studies such as Lane and Milesi-Ferretti (2008) or Portes and Rey (2005) who also underline the importance of goods trade in explaining the puzzling geography of international portfolio holdings via its potential value as an information variable.

5 Conclusion

I analyze a panel data set on country-level bilateral equity holdings as of the end-2006 to provide insights into the puzzling geographical patterns of international portfolio investments using a gravity model framework. Following Massa and Simonov's (2006) and Portes and Rey's (2005) inquiries on distinguishing behavioral-based familiarity effects from those originated by informational asymmetries in international financial markets, I argue that the preference revealed by investors towards physically closer geographies would undergo two different channels and examine the link via two quantitative measures as alternatives to geographical distance, i.e. *information distance* and *cultural distance*, respectively accounting for information- or familiarity-based aspects of investors' preference towards geographically close markets. While a generally recognized formula to provide a measure of the cultural distance between two countries is already available from previous studies, I believe that this paper is the first to propose and test a direct quantitative measure of the *information distance* between two countries. The results can be summarized as follows.

First and consistent with previous studies such as Papaioannou (2009), Lane and Milesi-Ferretti (2008) or Portes *et al.* (2001), regressions show that a gravity model of international trade in financial assets performs fairly well to explain the observed geographical patterns of bilateral equity holdings. This finding is robust to additional estimations including either the information or the cultural distance instead of the geographical distance, while variation in the data is better captured when geographical distance is used to substitute for transaction costs in international financial markets. The baseline gravity specification is typically suitable for the full sample and the developed-market source countries subsample, while most of the coefficients on distance proxies lack of explanatory power for the emerging-market source countries subsample. I also check for the model's accuracy using different estimation techniques including non-linear panel Tobit regressions. The results remain unchanged. Further, the impact of the physical distance is further pronounced for emerging-market countries subsample: When geographical distance is cut by 50%, emerging countries' foreign equity assets is more than doubled.

Second, beside the systematic effect of geographical distance on bilateral equity holdings, information and cultural distance measures also perform quite well within the baseline gravity model. Given the lack of correlation between these three distance measures, the proposed measures for information and cultural distances are likely to capture other aspects of the international market frictions along with physical distance. In particular, if the information and cultural distance between a pair of source and host countries decrease by 10%, bilateral equity assets held by the source countries' investors in the corresponding host are expected to increase by 5.4% and 3.5%, respectively. Hence, evidence suggests that the effect of the geography on international investment patterns is more likely to be a phenomenon linked to information asymmetries rather than reflecting the impact of cultural proximity between countries. Indeed, standardized coefficient estimates on information distance are systematically higher than those associated with the cultural distance regardless of the sample or the estimation technique used. When I employ these two distance measures jointly with geographical distance in 2SLS and IV estimations, the results from first-stage regressions show that information distance enters the models with expected sign and well-determined coefficients while the coefficient on the cultural distance generally lacks of statistical significance. This result also holds concerning non-linear Tobit regressions. Moreover, the impact of the geographical distance on the dependent variable is reduced in most cases once information distance is included into the model. In short, evidence supports Massa and Simonov's (2006) view that familiarity-driven investment is a response to information constraints as opposed to a behavioral heuristic (p. 634).

As a further check, I control for these results using additional gravity-type variables drawn from previous studies. I augment the baseline specification by adding several proxies to control for the effects of countries' economic development, openness to foreign trade, transparency and the degree of familiarity with the host market on portfolio investments. I also check if bilateral equity assets are driven by any diversification motive. Destination countries' economic development explains part of the dependent as suggested by significant estimates on the GDP per capita variable. Financial market sophistication and investor protection enter the regressions only when they are regressed separately from each other. Foreign openness is also helpful to

explain bilateral equity holdings given the significant estimation results concerning the mergers & acquisitions variable. The positive effect of host countries' transparency mainly goes through the corruptions perceptions index variable. This suggests that improving a country's transparency could enhance foreign investment inward and, consequently, erode some part of the home bias that foreign investors exhibit against the host country of interest. Finally, controlling for portfolio diversification, I find no evidence of any risk diversification motive in bilateral portfolios (Portes and Rey, 2005), since investors' tendency to hold stocks of a given country increases with the bilateral correlation coefficient between the source and host markets' returns (Aviat and Coeurdacier, 2007) and decrease with the host country's expected risk-adjusted return.

Beside the impact of the information distance, the most striking result concerns the systematic effect of bilateral trade on bilateral equity holdings. Specifically, bilateral trade always enters the regressions with positive and very-well determined coefficients regardless of the estimation technique or the sample employed. Previous studies such as Lane and Milesi-Ferretti (2008) or Aviat and Coeurdacier (2007) suggested that part of the effect of geographical distance on bilateral asset holdings could be related to bilateral trade relations between countries. In this paper, I check whether the effect of the trade on portfolio holdings undergoes an information- or culture-based story since these two alternative variables also appear to be good proxies for the geographical distance. Regressions highlight the relative importance of the information proximity against the cultural proximity in explaining the effect of bilateral trade in goods on the geographical patterns of international portfolio investments via its potential value as an information variable.

I consider the empirical work presented here as robust evidence that information asymmetries are central to understand the nature of international market frictions and the existing patterns of cross-border portfolio investments. The analysis also sheds light into the home bias puzzle suggesting that investors prefer to hold stocks in markets that are informationally close to their own home market. Although the puzzling preference for proximate geographies appears to be better explained by information proximity than cultural proximity, further work is needed to

effectively isolate the respective impacts of culture and information on the geographical distribution of cross-border asset portfolios and to investigate the true direction of causality running from culture to information exchange and *vice versa*. In fact, even if the two quantitative measures I propose for cultural and information distance are weakly correlated each other (and also with the geographical distance), it would be also useful to check for the extent to which culture and information are separable by means of additional instruments likely to capture other facets of the information exchange and cultural affinities between countries.

ACKNOWLEDGEMENTS

I thank Jonathan Hull from *Telegeography* and Florian Plank from *Euler Hermes* for their help to access data. Financial support from the *GREGOR* research team (Sorbonne Business School) is greatly acknowledged.

REFERENCES

Adler M. and B. Dumas (1983), "International portfolio choice and corporation finance: A synthesis", *Journal of Finance*, 38(3), p. 925-984

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

Aviat A. and N. Coeurdacier (2007), "The geography of trade in goods and asset holdings", *Journal of International Economics*, 71(1), p. 22-51

Baele L. and K. Inghelbrecht (2009), "Time-varying integration and international diversification strategies", *Journal of Empirical Finance*, 16(3), p. 368-387

Baele L., Pungulescu C. and J.T. Horst (2007), "Model uncertainty, financial market integration and the home bias puzzle", *Journal of International Money and Finance*, 26(4), p. 606-630

Barberis N.C. and R. Thaler (2003), "A survey of behavioral finance", in G.M. Constantinides, M. Harris and R.M. Stulz (eds.), Handbook of the Economics of Finance, vol. 1, chap. 18, p. 1053-1128

Barron J.M. and J. Ni (2008), "Endogenous asymmetric information and international equity home bias: The effects of portfolio size and information costs", *Journal of International Money and Finance*, 27(4), p. 617-635

Baxter M. and U.J. Jermann (1997), "The international diversification puzzle is worse than you think", *American Economic Review*, 87(1), p. 170-180

Bekaert G. and C.R. Harvey (2000), "Foreign speculators and emerging equity markets", *Journal of Finance*, 55(2), p. 565-613

Bekaert G. (1995), "Market integration and investment barriers in emerging equity markets", *World Bank Economic Review*, 9(1), p. 75-107

Beugelsdijk S. and B. Frijns (2010), "A cultural explanation of the foreign bias in international asset allocation", *Journal of Banking and Finance*, 34(9), p. 2121-2131

Black F. (1974), "International capital market equilibrium with investment barriers", *Journal of Financial Economics*, 1(4), p. 337-352

Brennan M.J. and H.H. Cao (1997), "International portfolio investment flows", *Journal of Finance*, 52(5), p. 1851-1880

Breusch T.S. and A.R. Pagan (1980), "The Lagrange multiplier test and its applications to model specifications in econometrics", *Review of Economic Studies*, 47(1), p. 239-253

Campbell R.A. and R. Kraussl (2007), "Revisiting the home bias puzzle: Downside equity risk", *Journal of International Money and Finance*, 26(7), p. 1239-1260

Chan K., Covrig V. and L. Ng (2005), "What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide", *Journal of Finance*, 60(3), p. 1495-1534

Chiou W.-J. P. (2009), "Benefits of international diversification with investment constraints: An over-time perspective", *Journal of Multinational Financial Management*, 19(2), p. 93-110

Cilibrasi R.L. and P.M.B. Vitanyi (2007), "The Google similarity distance", *IEEE Transactions on Knowledge and Data Engineering*, 19(3), p. 370-383

Cleaver C. and F.E. Warnock (2003), "Financial centers and the geography of capital flows", *International Finance*, 6(1), p. 27-59

Cooper I. and E. Kaplanis (1994), "Home bias in equity portfolios, inflation hedging, and

international capital market equilibrium", *Review of Financial Studies*, 7(1), p. 45-60

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

Dahlquist M., Pinkowitz L., Stulz R.M. and R. Williamson (2003), "Corporate governance and the home bias", *Journal of Financial and Quantitative Analysis*, 38(1), p. 87-110

De Jong E. and R. Semenov (2002), "Cross-country differences in stock market development: A cultural view", EFA 2002 Berlin meetings presented paper

Djankov S., Ganser T., McLiesh C., Ramalho R. and A. Shleifer (2008), "The effect of corporate taxes on investment and entrepreneurship", NBER working paper, no. 13756

Driessen J. and L. Laeven (2007), "International portfolio diversification benefits: Cross-country evidence from a local perspective", *Journal of Banking and Finance*, 31(6), p. 1693-1712

Dvorak T. (2005), "Do domestic investors have an information advantage? Evidence from Indonesia", *Journal of Finance*, 60(2), p. 817-839

Errunza V. and E. Losq (1985), "International asset-pricing under mild segmentation: theory and test", *Journal of Finance*, 40(1), p. 105-124

Faruqee H., Li S. and I.K. Yan (2004), "The determinants of international portfolio holdings and home bias", IMF working paper, no. 04/34

French K.R. and J.M. Poterba (1991), "Investor diversification and international equity markets", *American Economic Review*, 81(2), p. 222-226

Gehrig T. (1993), "An information based explanation of the domestic bias in international equity investment", *Scandinavian Journal of Economics*, 95(1), p. 97-109

Gelos R.G. and S.-J. Wei (2005), "Transparency and international portfolio holdings", *Journal of Finance*, 60(6), p. 2987-3020

Grinblatt M. and M. Keloharju (2001), "How distance, language, and culture influence stockholdings and trades?", *Journal of Finance*, 56(3), p. 1053-1073

Guiso L., Sapienza P. and L. Zingales (2003), "People's opium? Religion and economic attitudes", *Journal of Monetary Economics*, 50(1), p. 225-282

Hau H. and H. Rey (2008), "Home bias at the fund level", *American Economic Review*, 98(2), p. 333-338

Hau H. (2001), "Location matters: An examination of trading profits", *Journal of Finance*, 56(5), p. 1959-1983

Hausman J.A. (1978), "Specification tests in econometrics", *Econometrica*, 46(6), 1251-1271

Hofstede G. (1983), "National cultures in four dimensions", *International Journal of Management and Organization*, 13(2), p. 46-74

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

Kahneman D. and A. Tversky (1979), "Prospect theory: An analysis of decision under risk", *Econometrica*, 47(2), p. 263-291

Kang J.-K. and R.M. Stulz (1997), "Why is there a home bias? An analysis of foreign equity ownership in Japan", *Journal of Financial Economics*, 46(1), p. 3-28

Karlsson A. and A. Norden (2007), "Home sweet home: Home bias and international diversification among individual investors", *Journal of Banking and Finance*, 31(2), p. 317-333

Kogut B. and H. Singh (1988), "The effect of national culture on the choice of entry mode",

Journal of International Business Studies, 19(3), p. 411-432

La Porta R., Lopez-De-Silanes F., Shleifer A. and R.W. Vishny (2000), "Investor protection and corporate governance", *Journal of Financial Economics*, 58(1-2), p. 3-27

La Porta R., Lopez-De-Silanes F., Shleifer A. and R.W. Vishny (1997), "Legal determinants of external finance", *Journal of Finance*, 52(3), p. 1131-1150

Lane P.R. and G.M. Milesi-Ferretti (2008), "International investment patterns", *Review of Economics and Statistics*, 90(3), p. 538-549

Lane P.R. and G.M. Milesi-Ferretti (2004), "International investment patterns", IMF working paper, no. 04/134

Lee S.-H., Shenkar O. and J. Li (2008), "Cultural distance, investment flow, and control in cross-border cooperation", *Strategic Management Journal*, 29(10), p. 1117-1125

Lucas R.E. (1990), "Why doesn't capital flow from rich to poor countries?", *American Economic Review*, 80(2), p. 92-96

Lucey B.M. and Q. Zhang (2010), "Does cultural distance matter in international stock market comovements? Evidence from emerging economies around the world", *Emerging Markets Review*, 11(1), p. 62-78

Lütje T. and L. Menkhoff (2007), "What drives home bias? Evidence from fund managers' views", *International Journal of Finance and Economics*, 12(1), p. 21-35

Magi A. (2009), "Portfolio choice, behavioral preferences and equity home bias", *Quarterly Review of Economics and Finance*, 49(2), p. 501-520

Malloy C.J. (2005), "The geography of equity analysis", *Journal of Finance*, 60(2), p. 719-755

Martin P. and H. Rey (2004), "Financial super-markets: Size matters for asset trade", *Journal of International Economics*, 64(2), p. 335-361

Massa M. and A. Simonov (2006), "Hedging, familiarity and portfolio choice", *Review of Financial Studies*, 19(2), p. 633-685

Mishra A. and K. Daly (2006), "Where do Australians invest?", *Australian Economic Review*, 39(1), p. 47-59

Ni J. (2009), "The effects of portfolio size on international equity home bias puzzle", *International Review of Economics and Finance*, 18(3), p. 469-478

Nocetti D. (2006), "Markowitz meets Kahneman: Portfolio selection under divided attention", *Finance Research Letters*, 3(2), p. 106-113

Obstfeld M. and K. Rogoff (2000), "The six major puzzles in international macroeconomics: Is there a common cause?", NBER working paper, no. 7777

Papaioannou E. (2009), "What drives international financial flows? Politics, institutions and other determinants", *Journal of Development Economics*, 88(2), p. 269-281

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

Portes R., Rey H. and Y. Oh (2001), "Information and capital flows: The determinants of transactions in financial assets", *European Economic Review*, 45(4-6), p. 783-796

Prasad E.S., Rajan R.G. and A. Subramanian (2007), "Foreign capital and economic growth", NBER working paper, no. 13619

Reus T.H. and B.T. Lamont (2009), "The double-edged sword of cultural distance in international acquisitions", *Journal of International Business Studies*, 40(8), p. 1298-1316

Ricciardi V. (2008a), "The psychology of risk: The behavioral finance perspective", in F.J. Fabozzi

(ed.), Handbook of Finance, vol. II, p. 85-111

Ricciardi V. (2008b), "Risk: Traditional finance versus behavioral finance", in F.J. Fabozzi (ed.), Handbook of Finance, vol. III, p. 11-38

Rowland P.F. (1999), "Transaction costs and international portfolio diversification", *Journal of International Economics*, 49(1), p. 145-170

Sarkissian S. and M.J. Schill (2004), "The overseas listing decision: New evidence of proximity preference", *Review of Financial Studies*, 17(3), p. 769-809

Shleifer A. (2000), Inefficient Markets: An Introduction to Behavioral Finance, Oxford University Press, Oxford

Solnik B.H. (1974), "An equilibrium model of international capital market", *Journal of Economic Theory*, 8(4), p. 500-524

Sorensen B.E., Wu Y.-T., Yosha O. and Y. Zhu (2007), "Home bias and international risk sharing: Twin puzzles separated at birth", *Journal of International Money and Finance*, 26(4), p. 587-605

Stein E.H. and C. Daude (2007), "Longitude matters: Time zones and the location of FDI", *Journal of International Economics*, 71(1), p. 96-112

Stockman A.C. and H. Dellas (1989), "International portfolio nondiversification and exchange rate variability", *Journal of International Economics*, 26(3-4), p. 271-289

Strong N. and X. Xu (2003), "Understanding the equity home bias: Evidence from survey data", *Review of Economics and Statistics*, 85(2), p. 307-312

Stulz R.M. and R. Williamson (2003), "Culture, openness, and finance", *Journal of Financial Economics*, 70(3), p. 313-349

Stulz R.M. (1981), "On the effects of barriers to international investment", *Journal of Finance*, 36(4), p. 923-934

Suh J. (2005), "Home bias among institutional investors: A study of the Economist Quarterly Portfolio Poll", *Journal of the Japanese and International Economies*, 19(1), p. 72-95

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

Tinbergen J. (1962), Shaping the World Economy: Suggestions for an International Economic Policy", The Twentieth Century Fund, New York

Uppal R. and T. Wang (2003), "Model misspecification and underdiversification", *Journal of Finance*, 58(6), p. 2465-2486

Warnock F.E. (2002), "Home bias and high turnover reconsidered", *Journal of International Money and Finance*, 21(6), p. 795-805

Wei S.-J. (2000), "How taxing is corruption on international investors?", *Review of Economics and Statistics*, 82(1), p. 1-11

White H. (1980), "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity", *Econometrica*, 48(4), p. 817-838

APPENDIX. Data: Sample, definitions and sources

Reporting (source) countries in the CPIS 2006 files:

Argentina, Australia, Brazil, Canada, Czech Republic, Finland, France, Germany, Hungary, India, Indonesia, Italy, Japan, Malaysia, Mexico, Netherlands, Poland, South Africa, South Korea, Spain, Sweden, Turkey, United Kingdom, United States.

Destination countries in the CPIS 2006 files:

Albania, Algeria, Argentina, Australia, Austria, Azerbaijan, Bangladesh, Belarus, Belgium, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Cameroon, Canada, Chile, China, Colombia, Congo Republic, Croatia, Czech Republic, Denmark, Ecuador, Egypt, Estonia, Finland, France, Gabon, Georgia, Germany, Ghana, Greece, Guatemala, Guyana, Hungary, Iceland, India, Indonesia, Iran, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kuwait, Latvia, Liberia, Libya, Lithuania, Macedonia, Madagascar, Malawi, Malaysia, Mexico, Morocco, Namibia, Nepal, Netherlands, New Zealand, Nigeria, Norway, Oman, Pakistan, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saudi Arabia, Senegal, Serbia, Slovak Republic, Slovenia, South Africa, South Korea, Spain, Sri Lanka, Swaziland, Sweden, Taiwan, Tanzania, Thailand, Tunisia, Turkey, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Venezuela, Vietnam, Zambia, Zimbabwe.

Complement to variable definitions and data sources:

1) Economic development:

$GDPPC_j$: Country j's nominal GDP per capita. Source: IMF World Economic Outlook database. $FMSOP_j$: Country j's financial market sophistication index. Source: World Economic Forum's Global Competitiveness Index, www.weforum.org. $INVPRO_j$: Country j's investor protection index.

Source: The World Bank's "Doing Business Project" database.

2) Openness:

$FTRADE_j$: The foreign trade volume of country j's as a percentage of its GDP. Source: Author's own calculations based on trade statistics obtained from the web site www.trademap.org, GDP data is from the IMF's World Economic Outlook database. MA_j: The monetary sum of the country j's cross-border mergers and acquisitions scaled by its GDP. Source: Author's own calculations based on data covering the international mergers and acquisitions as reported by the World Investment Report (UNCTAD, 2007).

3) Familiarity:

$BTRADE_{ij}$: The ratio of the country i's bilateral trade volume with country j to the total foreign trade volume of country i. Source: Author's own calculations based on data from www.trademap.org. $LANG_{ij}$: A dummy variable equal to 1 if countries i and j share a common language, 0 otherwise. Sources: CIA World Factbook. $LEGAL_{ij}$: A dummy variable equal to 1 if countries i and j's legal systems come from the same origin, and 0 otherwise. Source: La Porta et al. (1997). REG_{ij}: A dummy variable equal to 1 if countries i and j are in the same geographical region, and 0 otherwise. Source: The World Bank.

4) Transparency:

CPI_j: The corruption perceptions index. Source: Transparency International, www.transparency.org. $JUDIN_j$: The judiciary system independence index. Source: The Economic Freedom Network, www.freetheworld.com. CMC_j: The capital market controls index. Source: The Economic Freedom Network, www.freetheworld.com.

5) Portfolio diversification:

RET_j: The expected return on the country j's market index. $SHRP_j$: The risk-adjusted return on country j's market index defined as the ratio of mean excess return to the standard deviation (for purpose of computational facility, the international risk-free rate is assumed to be equal to zero).

COR_{ij}: The correlation coefficient between the country i and j's market returns. Source: All variables are computed using index data (dividends included) over the 5-year length period running from January 2002 to December 2006. Whenever there is lack of a directly observable MSCI index for a particular country, I substitute it by a regional MSCI index; that is I employ i) the MSCI Emerging Asia index for Bangladesh and China, ii) the MSCI Emerging Markets Latin America index for Ecuador, Jamaica, Uruguay and Venezuela, iii) the MSCI Emerging Markets Eastern Europe index for Czech Republic, Bulgaria, Estonia, Romania and Slovak Republic, iv) the MSCI Emerging Markets index for Kenya, Nigeria and Zambia, and v) the MSCI Europe & Middle East index for Kuwait.

Table1. Hofstede's cultural dimensions and the countries' respective numerical scores

Country	C.	PDI	IDV	MAS	UAI	Country	C.	PDI	IDV	MAS	UAI
Argentina	1	49	46	56	86	Luxembourg [a]	1	40	60	50	70
Australia	2	36	90	61	51	Malaysia	4	104	26	50	36
Austria	1	11	55	79	70	Malta [a]	1	56	59	47	96
Bangladesh [a]	4	80	20	55	60	Mexico	3	81	30	69	82
Belgium	1	65	75	54	94	Morocco [a]	3	70	46	53	68
Brazil	3	69	38	49	76	Netherlands	2	38	80	14	53
Bulgaria [a]	3	70	30	40	85	New Zealand	2	22	79	58	49
Canada	2	39	80	52	48	Nigeria [d]	4	77	20	46	54
Chile	3	63	23	28	86	Norway	2	31	69	8	50
China [a]	4	80	20	66	30	Pakistan	3	55	14	50	70
Colombia	3	67	13	64	80	Panama	3	95	11	44	86
Costa Rica	3	35	15	21	86	Peru	3	64	16	42	87
Czech Rep. [a]	1	57	58	57	74	Philippines	4	94	32	64	44
Denmark	2	18	74	16	23	Poland [a]	1	68	60	64	93
Ecuador	3	78	8	63	67	Portugal	3	63	27	31	104
Egypt [b]	3	80	38	52	68	Romania [a]	3	90	30	42	90
El Salvador	4	66	19	40	94	Russia [a]	3	93	39	36	95
Estonia [a]	2	40	60	30	60	Saudi Arabia [b]	3	80	38	52	68
Ethiopia [c]	4	64	27	41	52	Sierra Leone [d]	4	77	20	46	54
Finland	2	33	63	26	59	Singapore	4	74	20	48	8
France	1	68	71	43	86	Slovakia [a]	4	104	52	110	51
Germany	1	35	67	66	65	South Africa	1	49	65	63	49
Ghana [d]	4	77	20	46	54	South Korea	3	60	18	39	85
Greece	3	60	35	57	112	Spain	1	57	51	42	86
Guatemala	3	95	6	37	101	Surinam [a]	3	85	47	37	92
Hong Kong	4	68	25	57	29	Sweden	2	31	71	5	29
Hungary [a]	1	46	80	88	82	Switzerland	1	34	68	70	58
India	4	77	48	56	40	Taiwan	3	58	17	45	69
Indonesia	4	78	14	46	48	Tanzania [c]	4	64	27	41	52
Iran	4	58	41	43	59	Thailand	3	64	20	34	64
Iraq [b]	3	80	38	52	68	Trinidad [a]	4	47	16	58	55
Ireland	2	28	70	68	35	Turkey	3	66	37	45	85
Israel	1	13	54	47	81	Unt. Arab Em. [b]	3	80	38	52	68
Italy	1	50	76	70	75	UK	2	35	89	66	35
Jamaica	4	45	39	68	13	USA	2	40	91	62	46
Japan	1	54	46	95	92	Uruguay	3	61	36	38	100
Kenya [c]	4	64	27	41	52	Vietnam [a]	4	70	20	40	30
Kuwait [b]	3	80	38	52	68	Venezuela	3	81	12	73	76
Lebanon [b]	3	80	38	52	68	Zambia [c]	4	64	27	41	52
Libya [b]	3	80	38	52	68						

Notes: a) estimated values, b) regional estimated values for "Arab World", c) regional estimated values for "East Africa", d) regional estimated values for "West Africa". The second column labeled "C." displays the country's membership score resulting from a four-means cluster using raw data from www.geert-hofstede.com.

Table 2. Descriptive statistics, raw data set

Variable		Obs.	Mean	St. dev.	Min.	Max.
Dependent						
Assets		1042	7000.63	37122.83	0.00	673978.00
Size and distance						
Size	(+)	2736	5.60e+11	2.76e+12	6.89e+07	5.78e+13
Distance, geographical	(–)	1626	7490.04	4674.48	84.00	19857.00
Distance, information	(–)	475	247.66	28.89	136.38	350.46
Distance, cultural	(–)	1560	268.42	104.26	28.05	692.68
Economic Development						
GDP per capita	(+)	2708	11331.09	15552.05	122.50	72768.10
Fin. market sophistication	(+)	2328	47.54	22.23	0.00	98.26
Investor protection	(+)	2664	48.90	16.46	0.00	100.00
Openness						
Foreign trade	(+)	2688	5.29	9.45	0.00	100.00
Mergers & Acquisitions	(+)	2160	2.85	3.24	0.00	14.04
Familiarity						
Bilateral trade	(+)	2700	0.81	2.85	0.00	68.93
Language	(+)	2856	0.47	0.49	0.00	1.00
Legal origin	(+)	945	0.32	0.46	0.00	1.00
Region	(+)	2856	0.26	0.44	0.00	1.00
Transparency						
Corruption perceptions	(+)	2664	31.04	27.91	2.56	100.00
Judiciary independence	(–)	2280	52.86	24.93	0.00	100.00
Capital market controls	(–)	2376	52.58	20.84	0.00	93.38
Portfolio Diversification						
Expected return	(+)	2040	2.02	1.09	–2.27	4.34
Sharpe ratio	(+)	2040	31.46	14.16	–23.58	64.29
Correlation	(–)	2040	50.21	22.72	–38.83	100.00

Notes: The table reports summary statistics computed from raw data. ASSETS are expressed in US$ millions and, GDPPC in US$ terms. Geographical distance is expressed in kilometers while the informational and cultural distances are without unit. The five sets of other explanatory variables are 1) Economic development variables, including the GDP per capita (GDPPC), financial market sophistication (FMSOP), and investor protection (INVPRO); 2) Openness variables, including foreign trade volume (FTRADE) and the mergers & acquisitions (MA) (both of them expressed as percentages of the country's GDP); 3) Familiarity variables, including the bilateral trade volume (BTRADE) as percentage of the source country's GDP, common language dummy (LANG), common legal origin dummy (LEG), common geographical region dummy (REG); 4) Transparency variables, including corruption perceptions index (CPI), judiciary system's independence index (JUDIN), and capital market controls index (CMC); and 5) Portfolio diversification variables, including the expected return (RET), Sharpe ratio (SHRP), and return correlations between the source and host countries (COR).

Table 3. The gravity model of bilateral equity holdings

Dependent: $\log(Assets_{ij})$	(1) GLS	(2) GLS	(3) GLS	(4) G2SLS	(5) G2SLS	(6) G2SLS
Panel A. All countries						
$Return_j$	–1.3818*** (–2.32)	–1.7605*** (–2.69)	–1.2727*** (–2.04)	–1.3637*** (–2.90)	–0.5209 (–0.86)	–0.9590** (–2.13)
$Size_{ij}$	1.2671*** (14.95)	1.1581*** (14.06)	1.1508*** (13.74)	1.3054*** (19.37)	1.4838*** (11.06)	1.3101*** (18.37)
$Geographical\ distance_{ij}$	–0.8865*** (–9.26)			–1.1632*** (–5.27)	–2.1981*** (–3.02)	–0.9959*** (–4.80)
$Information\ distance_{ij}$		–0.5017*** (–3.26)		[a] 0.4610*** (10.47)		[a] 0.4613*** (9.46)
$Cultural\ distance_{ij}$			–0.3280*** (–4.07)		[a] 0.1491*** (3.22)	[a] 0.0346 (0.81)
# of obs.	383	383	358	383	358	358
R^2 (overall)	0.4812	0.3882	0.3911	0.4791	0.4198	0.4827
Wald $\chi 2$	560.24	423.42	408.74	610.19	330.03	531.90
Panel B. Developed-market source countries						
$Return_j$	–1.7540* (–2.34)	–2.1585*** (–2.67)	–1.7383** (–2.17)	–1.5894*** (–3.48)	–0.0471 (–0.0474)	–1.2728*** (–2.96)
$Size_{ij}$	1.3683*** (15.86)	1.2720*** (14.13)	1.2950*** (12.79)	1.4001*** (22.06)	1.5991*** (10.92)	1.4160*** (21.14)
$Geographical\ distance_{ij}$	–0.7039*** (–6.82)			–0.8394*** (–4.05)	–2.8667** (–2.18)	
$Information\ distance_{ij}$		–0.5118*** (–3.16)		[a] 0.6336*** (9.21)		[a] 0.6607*** (8.62)
$Cultural\ distance_{ij}$			–0.3376*** (–3.63)		[a] 0.1106* (2.00)	[a] –0.0203 (–0.39)
# of obs.	284	284	261	284	261	261
R^2 (overall)	0.6616	0.6060	0.5861	0.6596	0.4152	0.6445
Wald $\chi 2$	589.21	515.03	439.06	730.79	222.04	643.52
Panel C. Emerging-market source countries						
$Return_j$	–2.8360* (–1.97)	–2.5713 (–1.59)	–1.2621 (–0.79)	–2.6418* (–1.69)	–0.9221 (–0.44)	–1.5147 (–0.96)
$Size_{ij}$	0.7927*** (3.90)	0.5305*** (2.93)	0.5831*** (3.32)	0.8087*** (3.21)	0.4929 (0.81)	0.7881*** (3.02)
$Geographical\ distance_{ij}$	–1.1792*** (–4.04)			–1.0300 (–1.25)	0.5688 (0.24)	–0.6220 (–0.80)
$Information\ distance_{ij}$		–0.2048 (–0.84)		[a] 0.2241*** (4.47)		[a] 0.2281*** (4.29)
$Cultural\ distance_{ij}$			0.0863 (0.31)		[a] 0.1282 (1.60)	[a] 0.0686 (0.92)
# of obs.	99	99	97	99	97	97
R^2 (overall)	0.1456	0.0795	0.0649	0.1394	0.0176	0.1157
Wald $\chi 2$	49.28	27.11	29.20	40.91	27.80	35.51

Notes: The table reports GLS and G2SLS regression results for the gravity model of international equity holdings. The dependent variable is log of equity assets held by the source country i in the host country j. $Return_j$ is lag–5 year average return on country j's stock market. $Size_{ij}$ is log of the source and host countries' GDP products. $Geographical\ distance_{ij}$, $Information\ distance_{ij}$, and $Cultural\ distance_{ij}$ are (log of the) the geographical, information and cultural distances between countries i and j. Concerning the G2SLS specification in columns (4) to (6), "a" denotes coefficient estimates from first-stage regressions. Heteroskedasticity consistent t-statistics obtained using the White (1980) method, are provided in parenthesis below the parameter estimates. Statistical significance at 10% (resp. 5 and 1%) is denoted by *** (resp. ** and *).

Table 4. The gravity model of bilateral equity holdings

Dependent: $\log(Assets_{ij})$	(1) Panel OLS	(2) Panel OLS	(3) Panel OLS	(4) IV	(5) IV	(6) IV
Panel A. All countries						
$Return_j$	–1.4516** (–2.16)	–1.8509** (–2.49)	–1.3222* (–1.88)	–1.3746*** (–2.85)	–0.6258 (–1.05)	–0.9751** (–2.13)
$Size_{ij}$	1.2704*** (14.26)	1.1662*** (13.08)	1.1569*** (12.36)	1.3078*** (18.95)	1.4789*** (11.42)	1.3146*** (18.11)
$Geographical\ distance_{ij}$	–0.8641*** (–8.99)			–1.1735*** (–5.21)	–2.1643*** (–3.05)	–1.0137*** (–4.83)
$Information\ distance_{ij}$		–0.5438*** (–4.11)		[a] 0.4634*** (10.22)		[a] 0.4654*** (9.26)
$Cultural\ distance_{ij}$			–0.3278*** (–4.40)		[a] 0.1514*** (3.20)	[a] 0.0345 (0.78)
# of obs.	383	383	358	383	358	358
R^2 (within)	0.6143	0.6025	0.5649	0.6328	0.4508	0.6314
F-statistic	20.96	27.19	21.60	21.42	16.29	24.28
Panel B. Developed-market source countries						
$Return_j$	–1.5050** (–2.05)	–1.8366** (–2.38)	–1.4938* (–1.92)	–1.4146*** (–3.11)	–0.0442 (–0.04)	–1.1382*** (–2.65)
$Size_{ij}$	1.4284*** (15.89)	1.3488*** (15.01)	1.3707*** (13.55)	1.4389*** (22.57)	1.5996*** (10.74)	1.4547*** (21.55)
$Geographical\ distance_{ij}$	–0.6743*** (–6.71)			–0.8704*** (–4.16)	–2.8662** (–2.12)	–0.7604*** (–3.95)
$Information\ distance_{ij}$		–0.5533*** (–3.70)		[a] 0.6356*** (9.16)		[a] 0.6687*** (8.60)
$Cultural\ distance_{ij}$			–0.3158*** (–3.76)		[a] 0.1102* (1.95)	[a] –0.0301 (–0.58)
# of obs.	284	284	261	284	261	261
R^2 (within)	0.7460	0.7206	0.6947	0.7424	0.1990	0.7404
F-statistic	12.49	11.72	11.43	11.24	4.23	12.33
Panel C. Emerging-market source countries						
$Return_j$	–2.6541* (–1.68)	–2.3481 (–1.39)	–0.8541 (–0.54)	–2.5639 (–1.61)	–0.7269 (–0.36)	–1.2991 (–0.82)
$Size_{ij}$	0.9018*** (4.33)	0.6459*** (3.30)	0.7202*** (3.77)	0.8653*** (3.44)	0.6411 (1.28)	0.8863*** (3.57)
$Geographical\ distance_{ij}$	–1.2362*** (–3.78)			–1.0781 (–1.33)	0.3318 (0.17)	–0.6850 (–0.95)
$Information\ distance_{ij}$		–0.2535 (–1.14)		[a] 0.2351*** (4.64)		[a] 0.2413** (4.63)
$Cultural\ distance_{ij}$			0.0495 (0.21)		[a] 0.1493* (1.86)	[a] 0.0903* (1.24)
# of obs.	99	97	97	99	97	97
R^2 (within)	0.3837	0.3122	0.3084	0.3823	0.2541	0.3767
F-statistic	6.58	7.55	8.33	7.95	7.43	8.86

Notes: The table reports panel OLS and instrumental variables regression results for the gravity model of international equity holdings. The dependent variable is log of equity assets held by the source country i in the host country j. $Return_j$ is lag–5 year average return on country j's stock market. $Size_{ij}$ is log of the source and host countries' GDP products. $Geographical\ distance_{ij}$, $Information\ distance_{ij}$, and $Cultural\ distance_{ij}$ are (log of the) the geographical, information and cultural distances between countries i and j. Concerning the IV specification in columns (4) to (6), "a" denotes coefficient estimates from first-stage regressions. F-statistic is designed to jointly test that all individual effects are zero, for which the null hypothesis is rejected at all conventional levels throughout the table. Heteroskedasticity consistent t-statistics obtained using the White (1980) method, are provided in parenthesis below the parameter estimates. Statistical significance at 10% (resp. 5 and 1%) is denoted by *** (resp. ** and *).

Table 5. The gravity model of bilateral equity holdings

Dependent: $\log(Assets_{ij} + \varepsilon)$	(1) Panel Tobit	(2) Panel Tobit	(3) Panel Tobit	(4) Panel Tobit	(5) Panel Tobit	(6) Panel Tobit
Panel A. All countries						
$Return_j$	1.0881***	1.2914***	1.2243***	–0.4876*	1.2538***	–0.4881*
	(3.17)	(3.35)	(3.30)	(–1.68)	(3.67)	(–1.68)
$Size_{ij}$	1.4654***	1.5599***	1.4393***	1.1250***	1.5273***	1.1248***
	(21.31)	(21.61)	(18.26)	(17.51)	(20.51)	(17.43)
$Geographical\ distance_{ij}$	–0.9544***			–0.5899***	–0.9302***	–0.5788***
	(–7.59)			(–5.61)	(–7.44)	(–5.61)
$Information\ distance_{ij}$		–0.5947***		–0.6414***		–0.6422***
		(–4.44)		(–6.25)		(–5.91)
$Cultural\ distance_{ij}$			–0.4471***		–0.2870***	0.0021
			(–3.85)		(–2.67)	(0.02)
# of obs.	421	362	398	362	395	362
R^2 (pseudo)	0.4996	0.3854	0.4034	0.4662	0.4932	0.4664
$\chi 2$	483.80	594.91	361.41	676.91	460.64	676.91
Panel B. Developed-market source countries						
$Return_j$	1.2097***	1.4296***	1.2998***	–1.6479***	1.3621***	–1.6355***
	(3.34)	(3.60)	(3.34)	(–4.06)	(3.85)	(–4.02)
$Size_{ij}$	1.5994***	1.7426***	1.6319***	1.1153***	1.6383***	1.1181***
	(20.22)	(20.44)	(16.88)	(14.70)	(18.55)	(14.68)
$Geographical\ distance_{ij}$	–0.8563***			–0.4981***	–0.8589***	–0.4998***
	(–6.39)			(–5.05)	(–6.52)	(–5.07)
$Information\ distance_{ij}$		–0.6867***		–0.6840***		–0.6678***
		(–3.52)		(–6.38)		(–5.85)
$Cultural\ distance_{ij}$			–0.4693***		–0.3436	–0.0353
			(–3.63)		(–2.91)	(–0.41)
# of obs.	297	262	275	262	273	262
R^2 (pseudo)	0.4934	0.3840	0.4049	0.4297	0.4873	0.4097
$\chi 2$	448.03	676.99	328.33	767.86	415.81	768.52
Panel C. Emerging-market source countries						
$Return_j$	0.4427	0.4450	0.5911	0.4895	0.5024	0.4751
	(0.93)	(0.9027)	(1.12)	(0.91)	(1.00)	(0.87)
$Size_{ij}$	1.0554***	0.9331***	0.9105***	0.9944***	1.1350	0.9753***
	(7.28)	(7.59)	(6.94)	(6.94)	(7.51)	(6.80)
$Geographical\ distance_{ij}$	–0.9136***			–0.9149***	–0.9629***	–0.9457**
	(–2.83)			(–2.72)	(–2.99)	(–2.83)
$Information\ distance_{ij}$		–0.2435		–0.0675		–0.1247
		(–1.37)		(–0.24)		(–0.45)
$Cultural\ distance_{ij}$			–0.2201		0.0984	0.2782
			(–0.09)		(0.42)	(1.25)
# of obs.	124	100	123	100	122	100
R^2 (pseudo)	0.4972	0.3972	0.4090	0.4951	0.4969	0.4841
$\chi 2$	57.02	44.40	52.82	54.37	62.91	56.71

Notes: The table reports panel Tobit regression results for the gravity model of international equity holdings. The dependent variable is log of equity assets held by the source country i in the host country j, normalized by the source and host countries' GDP products. $Return_j$ is lag–5 year average return on country j's stock market. $Size_{ij}$ is log of the source and host countries' GDP products. $Geographical\ distance_{ij}$, $Information\ distance_{ij}$, and $Cultural\ distance_{ij}$ are (log of the) the geographical, information and cultural distances between countries i and j. t-statistics are provided in parenthesis below the parameter estimates through columns. Pseudo R^2 is defined as the squared correlation between the predicted and observed values of the dependent variable. Statistical significance at 10% (resp. 5 and 1%) is denoted by *** (resp. ** and *).

Table 6. Panel regressions with control variables: Estimations using geographical distance

Dependent: $\log\left(\frac{Assets_{ij}}{GDP_i \times GDP_j}\right)$	(1) All	(2) Economic development	(3) Familiarity	(4) Foreign openness	(5) Transparency	(6) Portfolio diversification
$DGEO_{ij}$	–0.0228	–0.0498	–0.0395	–0.1235***	–0.0483	–0.0234
	(–0.34)	(–1.45)	(–0.74)	(–3.22)	(–1.37)	(–0.73)
$GDPPC_j$	0.1170	0.1980***				
	(1.52)	(4.91)				
$FMSOP_j$	–0.1390	–0.1369				
	(–0.92)	(–1.42)				
$INVPRO_j$	–0.0448	0.1334				
	(–0.38)	(1.10)				
$BTRADE_{ij}$	0.2506***		0.3369***			
	(4.45)		(7.29)			
$LANG_{ij}$	0.0388		0.0702			
	(0.51)		(1.13)			
LEG_{ij}	–0.0400		–0.0992**			
	(–0.86)		(–2.17)			
REG_{ij}	–0.1828		–0.2210			
	(–1.44)		(–1.99)			
$FTRADE_j$	0.0010			–0.1864***		
	(0.01)			(–3.02)		
MA_j	0.1360**			0.1747***		
	(2.23)			(3.43)		
CPI_j	–0.0655				0.3141***	
	(–0.38)				(4.09)	
$JUDIN_j$	0.1464*				–0.0127	
	(1.61)				(–0.16)	
CMC_j	–0.4485***				–0.1786	
	(–2.64)				(–1.46)	
$SHRP_j$	–0.3270***					–0.7081***
	(–2.68)					(–5.30)
COR_{ij}	0.3142					0.6390***
	(1.51)					(5.11)
# of obs.	293	371	293	369	371	368
R^2 (within)	0.3230	0.1585	0.2308	0.0731	0.1234	0.1441
F statistic	10.69	16.16	14.79	12.55	14.38	20.88

Notes: The table reports panel OLS regression results for the gravity model of international equity holdings using geographical distance, $DGEO_{ij}$. The dependent variable is log of equity assets held by the source country i in the host country j. Predictor variables are 1) economic development variables including GDP per capita ($GDPPC_j$), financial market sophistication ($FMSOP_j$), investor protection ($INVPRO_j$); 2) familiarity variables including bilateral trade ($BTRADE_{ij}$), common language dummy ($LANG_{ij}$), common legal origin dummy (LEG_{ij}), common geographical region dummy (REG_{ij}); 3) openness variables including foreign trade to GDP ratio ($FTRADE_j$), mergers & acquisitions to GDP ratio (MA_j); 4) transparency variables including corruption perceptions (CPI_j), judiciary independence ($JUDIN_j$), capital market controls (CMC_j); and 5) portfolio diversification variables including Sharpe ratio ($SHRP_j$), bilateral market correlations (COR_{ij}). All regressors, except dummies, are specified in log levels. Regressions control for the source country fixed-effects. F-statistic is designed to jointly test that all fixed-effects parameters are zero, for which the null hypothesis is rejected at all conventional levels throughout the table. Heteroskedasticity consistent t-statistics obtained using the White (1980) method, are provided in parenthesis below the parameter estimates. Statistical significance at 10% (resp. 5 and 1%) is denoted by *** (resp. ** and *).

Table 7. Panel regressions with control variables: Estimations using information distance

Dependent: $\log\left(\frac{Assets_{ij}}{GDP_i \times GDP_j}\right)$	(1) All	(2) Economic development	(3) Familiarity	(4) Foreign openness	(5) Transparency	(6) Portfolio diversification
$DINF_{ij}$	–0.0427	–0.1581***	0.0021	–0.1542***	–0.1472***	–0.1809***
	(–0.68)	(–3.11)	(0.03)	(–3.07)	(–2.86)	(–5.20)
$GDPPC_j$	0.1234	0.2163***				
	(1.59)	(5.58)				
$FMSOP_j$	–0.1354	–0.1152				
	(–0.88)	(–1.13)				
$INVPRO_j$	–0.0459	0.1543				
	(–0.40)	(1.35)				
$BTRADE_{ij}$	0.2404***		0.3500***			
	(4.09)		(7.03)			
$LANG_{ij}$	0.0368		0.0766			
	(0.47)		(1.23)			
LEG_{ij}	–0.0442		–0.0945**			
	(–0.93)		(–1.97)			
REG_{ij}	–0.1769*		–0.1718**			
	(–1.81)		(–2.01)			
$FTRADE_j$	0.0090			–0.1198**		
	(0.15)			(–2.21)		
MA_j	0.1402**			0.2238***		
	(2.32)			(5.06)		
CPI_j	–0.0623				0.3357***	
	(–0.36)				(4.75)	
$JUDIN_j$	0.1491*				0.0013	
	(1.67)				(0.01)	
CMC_j	–0.4331				–0.1475	
	(–2.74)				(–1.22)	
$SHRP_j$	–0.3197***					–0.6964***
	(–2.66)					(–5.34)
COR_{ij}	0.2954					0.6344***
	(1.50)					(5.45)
# of obs.	293	372	293	370	373	368
R^2 (adjusted)	0.2887	0.1939	0.2162	0.0835	0.1508	0.1411
F statistic	10.74	22.95	14.90	17.04	22.04	21.45

Notes: The table reports panel OLS regression results for the gravity model of international equity holdings using information distance, $DINF_{ij}$. The dependent variable is log of equity assets held by the source country i in the host country j. Predictor variables are 1) economic development variables including GDP per capita ($GDPPC_j$), financial market sophistication ($FMSOP_j$), investor protection ($INVPRO_j$); 2) familiarity variables including bilateral trade ($BTRADE_{ij}$), common language dummy ($LANG_{ij}$), common legal origin dummy (LEG_{ij}), common geographical region dummy (REG_{ij}); 3) openness variables including foreign trade to GDP ratio ($FTRADE_j$), mergers & acquisitions to GDP ratio (MA_j); 4) transparency variables including corruption perceptions (CPI_j), judiciary independence ($JUDIN_j$), capital market controls (CMC_j); and 5) portfolio diversification variables including Sharpe ratio ($SHRP_j$), bilateral market correlations (COR_{ij}). All regressors, except dummies, are specified in log levels. Regressions control for the source country fixed-effects. F-stat is designed to jointly test that all fixed-effects parameters are zero, for which the null hypothesis is rejected at all conventional levels throughout the table. Heteroskedasticity consistent t-statistics obtained using the White (1980) method, are provided in parenthesis below the parameter estimates. Statistical significance at 10% (resp. 5 and 1%) is denoted by *** (resp. ** and *).

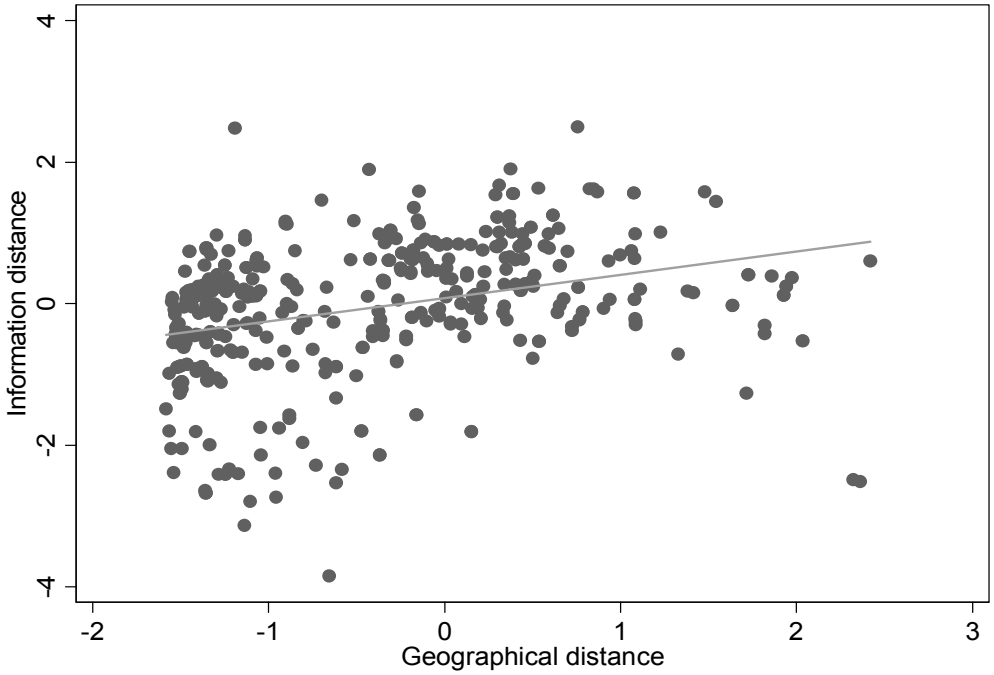
Table 8. Panel regressions with control variables: Estimations using cultural distance

Dependent: $\log\left(\frac{Assets_{ij}}{GDP_i \times GDP_j}\right)$	(1) All	(2) Economic development	(3) Familiarity	(4) Foreign openness	(5) Transparency	(6) Portfolio diversification
$DCULT_{ij}$	0.0064	–0.0198	–0.2250	–0.0465**	–0.0191	–0.0412**
	(0.24)	(–0.92)	(–0.82)	(–2.08)	(–0.84)	(–2.04)
$GDPPC_j$	0.1327	0.2056***				
	(1.48)	(5.03)				
$FMSOP_j$	–0.1215	–0.1012				
	(–0.81)	(–1.02)				
$INVPRO_j$	–0.0908	–0.0588				
	(–0.79)	(–0.55)				
$BTRADE_{ij}$	0.2594***		0.3447***			
	(4.75)		(7.49)			
$LANG_{ij}$	0.0444		0.0546			
	(0.50)		(0.76)			
LEG_{ij}	–0.0281		–0.1126**			
	(–0.53)		(–2.10)			
REG_{ij}	–0.1629*		–0.1829**			
	(–1.65)		(–2.21)			
$FTRADE_j$	–0.0123			–0.1044**		
	(–0.21)			(–2.04)		
MA_j	0.1296**			0.1707***		
	(2.15)			(3.42)		
CPI_j	–0.10852				0.3219***	
	(–0.58)				(4.26)	
$JUDIN_j$	0.1896**				0.0038	
	(2.18)				(0.06)	
CMC_j	–0.3409**				–0.2704**	
	(–2.06)				(–2.17)	
$SHRP_j$	–0.3021**					–0.5976***
	(–2.45)					(–4.63)
COR_{ij}	0.2466					0.6248***
	(0.92)					(4.80)
# of obs.	289	345	289	346	345	345
R^2 (within)	0.3227	0.1720	0.2334	0.0575	0.1410	0.1307
F statistic	10.29	15.23	14.15	9.71	13.01	16.65

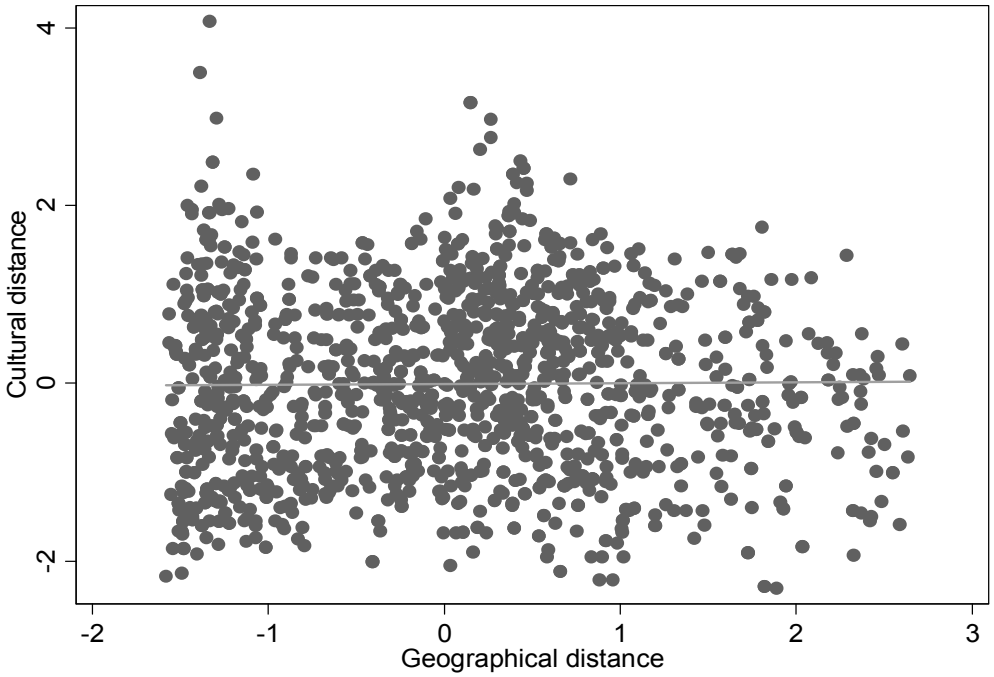
Notes: The table reports panel OLS regression results for the gravity model of international equity holdings using cultural distance, $DCULT_{ij}$. The dependent variable is log of equity assets held by the source country i in the host country j. Predictor variables are 1) economic development variables including GDP per capita ($GDPPC_j$), financial market sophistication ($FMSOP_j$), investor protection ($INVPRO_j$); 2) familiarity variables including bilateral trade ($BTRADE_{ij}$), common language dummy ($LANG_{ij}$), common legal origin dummy (LEG_{ij}), common geographical region dummy (REG_{ij}); 3) openness variables including foreign trade to GDP ratio ($FTRADE_j$), mergers & acquisitions to GDP ratio (MA_j); 4) transparency variables including corruption perceptions (CPI_j), judiciary independence ($JUDIN_j$), capital market controls (CMC_j); and 5) portfolio diversification variables including Sharpe ratio ($SHRP_j$), bilateral market correlations (COR_{ij}). All regressors, except dummies, are specified in log levels. Regressions control for the source country fixed-effects. F-stat is designed to jointly test that all fixed-effects parameters are zero, for which the null hypothesis is rejected at all conventional levels throughout the table. Heteroskedasticity consistent t-statistics obtained using the White (1980) method, are provided in parenthesis below the parameter estimates. Statistical significance at 10% (resp. 5 and 1%) is denoted by *** (resp. ** and *).

Figure 1. Information and cultural distance measures vs. geographical distance

Panel A. Information distance vs. Geographical distance



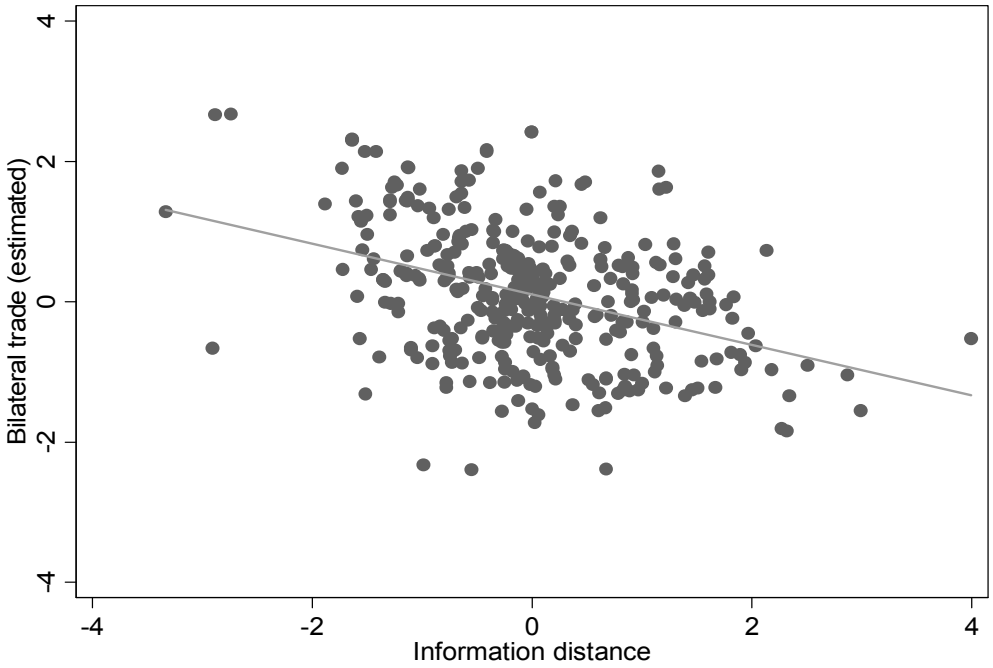
Panel B. Cultural distance vs. Geographical distance



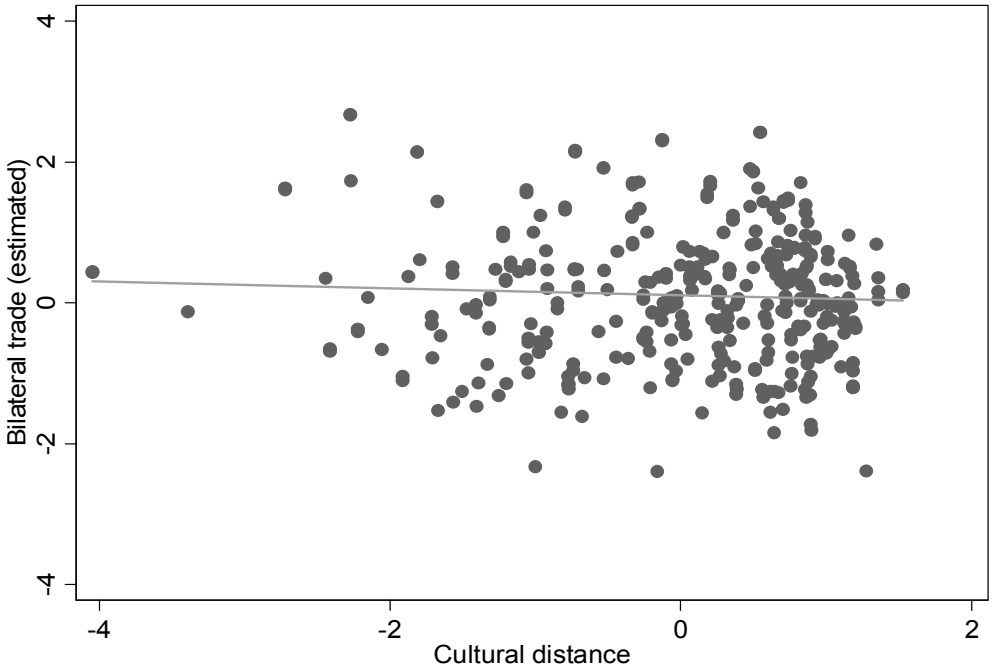
Notes: The figure shows the scatter plots of the informational distance and cultural distance measures against the geographical distance using standardized scores. Geographical distance is shown on the horizontal axis while the informational and cultural distance scores are shown on the vertical axis respectively in panels A and B.

Figure 2. The effect of bilateral trade on asset holdings: Information or Cultural distance?

Panel A. Bilateral trade vs. Information distance



Panel B. Bilateral trade vs. Cultural distance



Notes: The figure shows the scatter plots of bilateral trade (predicted values) against the information distance in panel A and cultural distance in panel B.