Spatial dependence in sovereign wealth funds' investments

Nicolas Debarsy*[1], Jean-Yves Gnabo†[2], and Malik Kerkour‡[2]

[1]CNRS, LEO, UMR 7322, Université d'Orléans, F-45067 Orléans, France.
[2]CeReFiM, University of Namur, Rempart de la Vierge 8, B-5000 Namur, Belgium.

December 7, 2015

Abstract

The aim of this paper is to identify the driving forces of Sovereign wealth funds' investments. For this, we develop an original econometric framework that quantifies the role of spatial dependence in the location of investments, and uses the Inverse Hyperbolic Sine transformation of the dependent variable in a spatial panel model context. This transformation copes with two features of net flows, namely an highly skewed distribution and the presence of zero and negative values. Using a large-scale database, we provide evidence of negative spatial dependence, investments in one country being on average at the expense of its neighbors.

JEL classification: C18; C21; C23; G15; G32; G38; F21
Keywords: Sovereign Wealth Funds; Pull Factors; Spatial Econometrics; Inverse Hyperbolic Sine Transformation; Interactions.

*Corresponding author. Email: nicolas.debarsy@univ-orleans.fr
†Email: jean-yves.gnabo@unamur.be
‡Email: malik.kerkour@unamur.be

"I don't want European citizens to wake up in several months' time and find that European companies belong to non-European capital, which bought at the share price's lowest point"

"This might be an opportunity to create our own sovereign wealth funds" (October 21, 2008, N. Sarkozy (President of France) at the European Parliament)

1 Introduction

Sovereign wealth funds (SWFs) are public investment agencies which manage part of the assets owned by national governments resulting from excess of exchange foreign reserves, oil or gas receipts as well as trade surpluses. Whether or not capital flows emanating from these state-owned entities are in the interest of the target country has been largely debated over the past decades. As illustrated in the above quote from the former French President, Nicolas Sarkozy, there have been long lasting fears and reluctance among the general public and policymakers regarding the arrival of SWFs in local markets mainly due to their lack of transparency as well as to the belief that those investments were designed to exert political influence on target firms or countries and to access foreign technology. Today, however, SWFs constitute a major source of capital for world economies with assets under management amounting to \$7.243 trillion,[1] leading most governments as noted by Megginson and Fotak (2014) to "court" SWF investments.

Given their increasing importance in advanced and emerging economies, the question of how countries can attract capital flows emanating from SWFs has become of major importance. In this paper, we propose to identify both country-level determinants and spatial interaction effects of cross-border SWF net flows. To this end, we develop an original econometric framework that allows (i) to explicitly model and to test the existence of spatial dependence in investment location and (ii) to accommodate the presence of skewed distributions of net investment flows due to extreme values as well as zero and negative net flows. Our procedure is then applied to an original large-scale database containing SWF's net flows to 43 countries over the 2004-2009 period.

So far, the research community has shown increasing interest in trying to understand the determinants of SWFs' investments at the firm level or at the country level. In the former case, the literature builts on existing evidence in corporate finance which has shown that firm-level conditions, such as firm-size (market capitalization, sales), firm-specific risk indicators (leverage, cash of firm, degree of financial constraint, analyst coverage, turnover) as well as firm-performance (ROA, ROE, CAPEX, stock market return, dividend yield) are the main drivers of firm attractiveness for capitals (Kotter and Lel, 2011; Fernandes, 2011; Avendano, 2010). With respect to these three criteria, Kotter and Lel (2011), Karolyi and Liao (2010) and Avendano (2010) argue that SWFs tend to invest in large and poorly performing companies which are financially and cash constrained. Beyond firm characteristics,

[1] SWFs institute website (*www.swfinstitute.org/sovereignwealthmap.html*), assets under management in September 2015.

macroeconomic factors along with the financial and institutional factors of the target country have also been documented as important in the decision process of SWF (see Karolyi and Liao (2010) and Chhaochharia and Laeven (2008) among others). Hence, Knill et al. (2012) find that the decision to invest depends positively on the correlation between the market return on the SWF and that on the target country's national market index, mitigating the role of risk diversification in the decision of investment. In the same vein, Fotak et al. (2008) find that SWFs investments tend to be concentrated, especially in the financial sector. Separating the decision to invest from the amount invested, Knill et al. (2012) conclude that the economic and financial variables are important factors for explaining the former but matter less in determining the size of investment.

One important dimension of SWFs decision which is absent from the above discussion pertains to the international allocation of capital. Is there evidence of domestic or foreign equity bias? Do cross-border investments concentrate in specific regions? And if so, why? These are examples of questions related to the spatial dimension of the portfolio allocation which have been discussed in the literature on SWF without being fully addressed (Megginson et al., 2013). In what follows, we review key elements of those discussions along with selected arguments regarding the specific role of interactions in explaining capital investments flow among countries.

The primary remarkable feature regarding the spatial allocation of SWFs' capitals lies in the large share of foreign assets hold in their portfolio. Looking at investments made by 15 SWFs over the period 1985-2011, Bortolotti et al. (2013) and Megginson et al. (2013) report that nearly 70% percent are channeled outside the home country. These figures starkly contrast with those usually reported for other categories of funds such as mutual funds or pension funds which exhibit very strong home equity bias (Bortolotti et al., 2013; Megginson and Fotak, 2014). At least two explanations can help understand this feature. First, conversely to other types of funds, one of the primary missions of SWF is to help its home country stabilize its wealth. Accordingly, cross-border investments are used as a diversification device which mitigates the impact of domestic economic downturns on the national wealth. Second, large demand in domestic assets could end up building financial bubbles in local markets. This risk is particularly important for countries hosting SWFs as their economies are often of small size compared to the amount under management.

Another interesting feature related to the spatial dimension pertains to the concentration of SWF investments in specific regions, notably in western economies which benefit from the lion's share of international SWF capitals with more than 50 % of them (see Megginson and Fotak (2014) and the Sovereign wealth fund institute). The rest is mainly invested in emerging market economies with specific regions such as Asia-Pacific receiving large flows and others such as Latin America being nearly absent from the map. Traditional arguments from the literature on international asset allocation to explain this feature relate to restrictions on international capital flows, institutional barriers or transaction costs. For instance, high transaction costs prevent from widespread investments of capital across countries and in turn from optimal diversification of risk. Regional concentration can also stem from the existence of privileged relationship between the home and the target country. For instance, Knill et al. (2012) and Chhaochharia and Laeven (2008) show that SWF tend to invest in countries with prevailing trade links. This feature is consistent with former results from the literature on foreign direct investment and international asset allocation which emphasize the critical role of symmetric information (cost) and familiarity in explaining the spatial allocation of international investments (Beugelsdijk and Frijns, 2010). Likewise, Hau (2001) shows that investors overweight assets they know better in their portfolio such as those issued by firm located in a country they are more familiar with or by multinationals. Those arguments have been more recently discussed within the frame of the behavioral finance theory in which investors are assumed to be boundedly rational. Accordingly, they consider foreign markets as more risky than they truly are, if they are located far away (Solnik, 2008), concentrating their investments in regions in which they have either information advantage or perceived familiarity. In a related context, Portes and Rey (2005) highlight the strong impact of the distance on cross-border equity flows. To explain this result, they argue that informational asymmetries lead to higher transaction costs between distant economies.[2] Aviat and Coeurdacier (2007) point out that the very large impact of distance on asset holdings in Portes and Rey (2005) is the consequence of the complementarity between trade in goods and trade in assets. As such, they show that geographical distance, understood as transportation costs, affects asset holdings mainly through its impact on trade in goods. Overall, various proxies have been proposed in the literature to measure the level of familiarity between investors and target firms or economies such as language, trade, culture and geographic distance . These types of mechanisms have been well documented in the literature on foreign direct investments that specifically concentrates on conditions of multinational enterprise (MNE) choice of location for their production plan, with factors at play such as market access, market size or growth prospect.

Spatial dependence can also simply stem from well-documented economic and financial interactions across world economies. As such, countries receiving capitals from SWFs are open economies which trade with other partner countries or investing themselves in other countries. Those economic and financial relationships create dependence across countries that make capital flows from SWFs in a given country to be complementary and then to depend on what its neighbors receive.

The above discussion suggests the existence of spatial dependence in aggregate flow of investments received by host countries. Whether this dependence is playing a significant role in practice is not clear cut. Neither is the actual sign of the correlation as two opposite effects can be at play. Hence, if the choice to invest in a country is primarily based upon information or familiarity criteria, we do expect a positive sign: countries sharing similar characteristics should attract SWFs' capital more or less to the same extend. To illustrate this relationship, consider a SWF deciding to incur search and informational costs for investing in a specific country. By doing so, the fund implicitly becomes more informed about countries having a similar profile to the target country, thereby being also more likely to invest in these countries. Increased investments in one country should therefore be positively correlated with

[2]Separating out cultural proximities from pure informational symmetries remains a challenge for the empirical literature (Portes and Rey, 2005).

investments in its neighborhood (viewed as countries sharing similarities). Still, the sign of the relationship holds as long as countries are not set in competition. If similar economies are set in competition the sign is reversed and the expected correlation should be negative due to a continuous reallocation or re-balancing mechanism over time: an increase in the investment done in one country at time t comes at the expense of a similar country.

Although the literature on SWF determinants has been important over the past decade, empirical studies have been silent on the specific role of spatial interactions between host countries to explain the capitals' allocation. The present paper aims to fill the gap by assessing the effect of domestic determinants and spatial interactions on sovereign funds capitals' net flows. To this end, we adopt the perspective of the host country. As such, we can identify the factors leading an economy to attract more or less capitals emanating from SWFs while being agnostic upon their origin. This approach is in line with a large strand of the literature on foreign direct investment (Yu and Walsh, 2010; Kristjánsdóttir, 2005; Bénassy-Quéré et al., 2007) but departs from the traditional bilateral approach often adopted in the SWF literature (Knill et al., 2012; Megginson et al., 2013). Hence, our contribution to the existing literature is twofold. First, our approach is designed to explicitly take into account interactions in SWF's investment flows. As discussed above, there are ample anecdotal and more formal evidence from both the academic literature and professional commentators that SWF investments concentrate on specific geographic zones.[3] Accordingly, we conjecture that the amount flowing in and out a country is not independent from the amount of capital attracted by neighboring countries, the neighborhood being alternatively defined by geographic, cultural or economic proximity in this paper. To formally test the existence of spatial interactions in SWFs investment net flows, we propose to rely on the recent developments of spatial panel data models (Lee and Yu, 2010a,b, 2012). To the best of our knowledge, this approach has never been used to analyze SWFs' investments. The second contribution lies in the treatment proposed to deal with the so-called problem of skewness in the distribution of capital flows (Zhang, 2014, p.136). More specifically, the amount invested every year in target countries exhibits extreme values but also null and negative values. To deal with the problem of extreme values, the literature usually proposes to apply a log-transformation (Bénassy-Quéré et al., 2007). This step has the advantage to dampen the dispersion in the data, and especially the outliers, but cannot deal with zero and negative values. In this paper, we propose to tackle this issue by using the Inverse Hyperbolic Sine transformation of the dependent variable, developed by Johnson (1949) and applied to econometrics by Burbidge et al. (1988), which allows, in addition to reduce the skewness of the distribution and the impact of outliers, to take care of possible zeros and negative values in the dependent variable. To the best of our knowledge, this transformation has never been applied to spatial econometrics models. We contribute therefore to the literature by developing the estimation procedure accordingly. Specifically, we derive the likelihood function that accommodates the transformation and compute the associated matrices of impacts of a change in a regressor on the dependent variable.

The reminder of the paper is as follows: In Section 2, we present the spatial

[3] see www.swfinstitute.org/sovereignwealthmap.html

autoregressive model with the IHS transformation. Section 3 provides some details regarding the data and Section 4 reports our empirical findings. Finally, Section 5 concludes.

2 Econometric methodology

In this contribution, the dependent variable is defined as the net flows of investments from sovereigns funds in a given country and is highly skewed to the right (see Figure 1 and information given in section 3.1). To solve this problem of outliers which can severely bias the estimated parameters of the model, the traditional approach consists in log-transforming the variable so that the distribution of the transformed variable is closer to a normal distribution. However, in our case, the dependent variable also takes on null and negative values. Different solutions have been proposed in the literature to deal with the limits of standard log-transformation. Hence, in the presence of positive and null values of the dependent variable (presence of zero flows for instance), we can simply disregard the concerned observations. However, this approach may result in substantial loss of information (Aviat and Coeurdacier, 2007). An alternative solution consists in adding 1 to the concerned observations and include a dummy in the model which takes the value 1 for the modified values of the dependent variable (for further details, see Disdier and Head, 2008). However, the zero-flow dummy variable could be endogenous since it depends on the value taken by the dependent variable. Alternatively, Raballand (2003) proposes to replace zero values by a small constant, so that the model can be log-linearized and estimated without getting rid of the zero-flows. Nevertheless, the inserted value is arbitrary and does not reflect the reality (Linders and De Groot, 2006). Besides, Frankel et al. (1997) indicates that this transformation biases the least squares estimator. Another solution to deal with zero flows would be to use a Tobit model (Eaton and Tamura, 1994). However, this approach might not be suited for our case since in addition to zero investments, we also face negative values of net flows. Hence, our dependent variable cannot be considered as left censored.

In this contribution, we rely on a different approach which consists in using an alternative transformation to the logarithm. We apply the Inverse Hyperbolic Sine (IHS) transformation to the dependent variable. This IHS transformation, developed by Johnson (1949) and applied to econometrics by Burbidge et al. (1988) has been used among others in the context of Engel curve by Reynolds and Shonkwiler (1991); Yen and Jones (1997), to assess the impact of tax incentives on savings by Pence (2006) and to studies aimed at explaining wealth (Kennickell and Sunden, 1997; Carroll et al., 2003; Kapteyn and Panis, 2003). One of the advantages of the IHS transformation is that it allows the dependent variable to take on both positive and null values but also accommodates negative values while still dampening the outliers and is known to better handle extreme values than the Box-Cox transformation (Burbidge et al., 1988). [4] The IHS transformation applied to a variable y is presented

[4] Bickel and Doksum (1981) extend the Box-Cox transformation so that it can also handle negative value, but this transformation remains undefined when the variable takes on zero values. Also, John and Draper (1980) propose the "modulus transformation", another extension of the Box-Cox transformation, to address negative and zero values. However, this modulus transformation

in equation 1.

$$\omega(y) = \frac{\ln(\theta_0 y + (\theta_0^2 y^2 + 1)^{0.5})}{\theta_0} = \sinh^{-1}(\theta_0 y)/\theta_0. \tag{1}$$

This transformation is defined for all values of the scaling parameter θ_0. Besides, as it is symmetric around 0, we consider only $\theta_0 \geq 0$. Also, for large values of y, the IHS transformation corresponds to a vertical displacement of the logarithm: $\ln(\theta_0 y + (\theta_0^2 y^2 + 1)^{0.5}) \approx \ln 2\theta_0 + \ln y$. It can thus be interpreted in exactly the same way as a standard logarithmic dependent variable but unlike the logarithm, the IHS transformation is defined at zero and for negative values. As θ_0 approaches 0, the IHS transformation is linear for a larger proportion of its domain while it approximates the logarithmic transformation in a larger proportion of its domain when θ_0 is larger (Pence, 2006). Finally, we note that $\partial\omega(y)/\partial y = \frac{1}{\sqrt{1+\theta_0^2 y^2}}$.

In this paper, we estimate a spatial panel data model with random effects. Our procedure uses the approach developed by Mundlak (1978) and extended to spatial panel models by Debarsy (2012) to correct for the possible correlation between individual effects and regressors.

The random effects specification is presented in equation 2.

$$\begin{aligned} \omega(Y)_{nt} &= \lambda_0 W_n \omega(Y)_{nt} + X_{nt}\beta_0 + U_{nt} \quad t = 1, \cdots, T \\ U_{nt} &= c_{n0} + V_{nt} \end{aligned} \tag{2}$$

where Y_{nt} is the vector of the dependent variable for the n individuals at period t and X_{nt} is the matrix of exogenous regressors at period t of dimension $n \times K$. The error term U_{nt} is composed of 2 terms: $\mathbf{c}_{n0}$, the vector of individual effects and $V_{nt} = (v_{1t}, \cdots, v_{nt})$, the vector of idiosyncratic errors. We further have that W_n is a square matrix of dimension n modeling interactions between observations. Its construction is discussed in section 3. Finally, β_0 is the vector of dimension $k \times 1$ of unknown parameters to be estimated and λ_0 is the unknown coefficient measuring the intensity of interactions between observations, to be estimated.

Stacking model 2 for all periods, we obtain:

$$\begin{aligned} \omega(Y)_{nT} &= \lambda_0 (I_T \otimes W_n)\omega(Y)_{nT} + X_{nT}\beta_0 + U_{nT} \\ U_{nT} &= Z_c c_{n0} + V_{nT} \end{aligned} \tag{3}$$

with $Z_c = \iota_T \otimes I_N$. The variance-covariance matrix of U_{nT} is

$$\Omega_{nT} = (\iota_T \iota_T' \otimes I_N)\sigma_c^2 + \sigma^2 I_T \otimes I_N \tag{4}$$

Using the results of Wansbeek and Kapteyn (1982), we can rewrite the var-cov matrix in 4 as follows:

$$\Omega_{nT} = (T\sigma_\alpha^2 + \sigma^2)(\frac{1}{T}\iota_T \iota_T' \otimes I_N) + \sigma^2(\frac{1}{T}\iota_T \iota_T' \otimes I_N) \tag{5}$$

is not scale invariant, to the contrary of the IHS transformation used here.

where $E_T = I_T - \frac{1}{T}\iota_t\iota_T'$ is the demeaning operator. By writing $\Omega_{nT} = \sigma^2\Sigma$, with

$$\Sigma = \phi^{-1}(\frac{1}{T}\iota_T\iota_T' \otimes I_N) + (\frac{1}{T}\iota_T\iota_T' \otimes I_N)$$

where $\phi = \frac{\sigma^2}{(T\sigma_c^2+\sigma^2)}$. We can use Breusch (1987) results and compute $\Sigma^{-1} = \phi P_W + Q_W$ and $|\Sigma| = \phi^{-N}$.

Asymptotic properties of this model have been developed by Lee and Yu (2012) and the assumptions underlying these asymptotic properties for the QML estimator are reproduced below for the sake of clarity.

Assumption 1. *W_n is a non-stochastic spatial weights matrix with zero diagonals*

Assumption 2. *The disturbances $\{v_{it}\}$, $i = 1, \cdots, n$ and $t = 1, \cdots, T$ are i.i.d. across i and t and normally distributed with zero mean, variance σ_0^2 and $\mathcal{E}|v_{it}|^{4+\eta}$ for some $\eta > 0$.*

Assumption 3. *$S_n(\lambda) = I_n - \lambda W_n$ is invertible for all $\lambda \in \Lambda$, where Λ is a compact interval. Furthermore, λ_0 is in the interior of Λ.*

Assumption 4. *The elements of X_{nt} are non-stochastic and bounded, uniformly in n and t. Also, under the asymptotic setting in assumption 6, the limit of $\frac{1}{nT}\sum_{t=1}^{T}\widetilde{X}'\widetilde{X}$ exists and is nonsingular.*

Assumption 5. *W_n is uniformly bounded in both row and column sums in absolute value (for short, UB). Also, $S_n(\lambda)^{-1}$ is UB, uniformly in $\lambda \in \Lambda$.*

Assumption 6. *n is large, where T can be finite or large.*

Assumption 7. *$\mathbf{c}_{n0} \sim N(0, \sigma_c^2 I_n)$ and V_{nt} are i.i.d. and independent of X_{nT}. Also, $\mathbf{c}_{n0}$ is independent of V_{nT}.*

Assumption 1 is a normalization assumption which simply states that an observation does not interact with itself. Assumption 2 states the regularity conditions for the error term. It imposes homoskedasticity of the error term, a necessary condition for the QMLE to be consistent (see Lin and Lee, 2010). The invertibility condition in Assumption 3 guarantees the existence of equation 2. Besides, Assumption 3 imposes the parameter space Λ to be a compact set.

Assumption 4 states that if explanatory variables are included in the model, they should be uniformly bounded. Finally, Assumption 5 guarantees that spatial autocorrelation is limited to a manageable degree (Lee and Yu, 2010a) while Assumption 7 is needed to avoid correlation between independent variables and the error term. Defining $\gamma = (\tau,\ \lambda,\ \beta',\ \phi, \sigma^2)'$ and $\eta = (\tau,\ \lambda,\ \beta',)'$, the log-likelihood function associated to model 2 is written as:

$$\ln L(\gamma) = -\frac{NT}{2}\ln(2\pi) - \frac{NT}{2}\ln(\sigma^2) + \frac{N}{2}\ln\phi + T\ln|I_n - \lambda W_n| - \frac{1}{2}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\ln(\theta^2 y_{it}^2 + 1)\right] - \frac{1}{2\sigma^2}U_{nT}'(\eta)\Sigma^{-1}U_{nT}(\eta) \quad (6)$$

where $U_{nT}(\eta) = \omega(Y) - \lambda(I_T \otimes W)\omega(Y) - X\beta$.

Also, $T \ln |S_n(\lambda)| - \frac{1}{2}\sum_{i=1}^{n}\sum_{t=1}^{T} \ln(\theta^2 y_{it}^2 + 1)$ is the Jacobian of the transformation, composed of two terms. $T \ln |S_n(\lambda)|$ refers to the classical term coming from the SAR specification while $-\frac{1}{2}\sum_{i=1}^{n}\sum_{t=1}^{T} \ln(\theta^2 y_{it}^2 + 1)$ is the derivative of the transformed dependent variable with respect to the original dependent variable.

According to Mundlak (1978), the random effects model is a misspecified version of the fixed effects model because it ignores the possible correlation between individual effects and regressors. By controlling for this correlation, he shows that the coefficients of the random effects specification are identical to those of the fixed effects model. He thus proposes to set an auxiliary regression that will capture this possible correlation. As individual effects are time invariant, they should be correlated with the permanent component of the regressors, namely their time-average.

Hence, Mundlak (1978) proposes to augment the random effect model with the time-mean of independent variable to control for the possible correlation between individual effects and regressors. Debarsy (2012) extends this approach to spatial panel data models.[5] Adapting this approach to our model, we get:

$$\begin{aligned} \omega(Y)_{nt} &= \lambda_0 W_n \omega(Y)_{nt} + X_{nt}\beta_0 + \overline{\widetilde{X}}_n \pi_0 + U_{nt} \quad t = 1, \cdots, T \\ U_{nt} &= \alpha_{n0} + V_{nt} \end{aligned} \tag{7}$$

where $\widetilde{X}$ is the matrix of independent variables without the constant term, $\overline{\widetilde{X}}_n = \frac{1}{T}\sum_{t=1}^{T} \widetilde{X}_{nt}$ is the time average of $\widetilde{X}$ for all individuals and $\alpha_{n0} \sim IN(0, \sigma_\alpha^2)$.

The presence of correlation between individual effects and independent variables can be assessed by testing whether the hypothesis $\pi = 0$ can be rejected, using for instance a Wald or likelihod ratio test.

3 Data and empirical specification

The regression concerns cross-border SWFs net flows for 43 host countries from 2004 to 2009. The econometric model to be estimated is a spatial panel data framework with country random effects and the correction of Mundlak and is presented in equation 7. Y_{nt} is the vector of SWFs' cross-border net flows received by the 43 host countries of the sample at period t and $\omega(Y)_{nt}$ is the IHS transformed dependent variable.[6]

One of the main challenge given to the research community on SWF is the lack of official data. Despite the slight improvement over the years, SWFs as most financial investors have remained extremely reluctant to disclose a clear and comprehensive information about their portfolio allocation. In this paper, we construct a novel database on SWF net flows .

[5] As spatial panel data models are nonlinear, the estimators obtained using the Mundlak approach are not equivalent anymore to those obtained with the fixed effects approach.

[6] The considered countries are: Argentina, Australia, Austria, Bahrain, Belgium, Brazil, Canada, Chile, China, Colombia, Czech Republic, Denmark, Egypt, Finland, France, Germany, Greece, Hong Kong, Hungary, Indonesia, Ireland, Israel, Italy, Jordan, Malaysia, Mexico, Morocco, Netherlands, New Zealand, Norway, Peru, Philippines, Poland, Portugal, Russia, Singapore, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom, United States.

For this, we follow Avendano (2010) and Fernandes (2011) who extract their information from the FactSet database. FactSet is a major information source for institutional ownership. The construction of the dependent variable is described in subsection 3.1. The set of considered determinants is described in subsection 3.2. Finally, the interaction matrices used are described in subsection 3.3.

3.1 Dependent variable

We consider in our dependent variable equity cross-border SWFs net flows amount, excluding investments and divestments made by local SWF in their own country. Amount are expressed in millions of U.S. dollars. The sample consists of 33,927 (resp. 16,464) cross border investments (resp. divestments) in public firms made by the 24 largest SWFs. The total foreign acquisition (resp. sell) flows represents 1,157 (resp. 379) billions of U.S. dollars. We are careful to extract investment data for both the SWFs and their subsidiaries, which we define as entities in which SWF has at least a 50% ownership stake, following the literature (Megginson et al., 2013; Chhaochharia and Laeven, 2008; Avendano, 2010; Bernstein et al., 2009).

The 24 largest SWFs included in our sample hold in 2009 approximatively $5.93 trillion[7], representing 84% of total SWF holdings. We restrict our sample to the 2004-2009 period for the sake of homogeneity of our sample as the economic and financial turbulences observed in the aftermath of the Lehman brothers collapse and during the European sovereign debt crisis have dramatically changed the financial environment and the behavior of financial investors. In this respect, our analysis aims to shed light on the driving forces of SWF net flows in normal time.

We transform monthly holding data into share flows (acquisitions or sells) by computing the difference between the number of positions hold by the SWF in a specific company at time t and at time $t-1$. Then, we multiply this quantity by the average stock market price of the target firm for the associated month in order to obtain the amount invested or divested. [8] Eventually, we aggregate this information across firms, months and SWFs for retrieving the annual net investments per country which is designed to capture the attractiveness of the host countries. Descriptive statistics of our dependent variable are presented in Table 1. At the host country level, we observe in Table 1 both negative and positive net flows amounts, which justifies the use of the IHS transformation.

3.2 Explanatory variables

In this section, we detail the construction of the different explanatory variables used to estimate the model as well as the expected effect on investments' net flows. For the sake of clarity, the set of variables is broken down into two groups: (i) those related to the economic development and (ii) those related to the financial development.

[7]SWF institute website

[8] We consider the monthly average price instead of the one observed the day of the deal to mitigate the impact of the acquisition or divestment on the stock market price (Kotter and Lel, 2011; Dewenter et al., 2010; Fotak et al., 2008; Raymond, 2010; Fernandes, 2011; Bortolotti et al., 2013; Hesse and Sun, 2009).

Economic development is described as a sustainable increase in living standards of a certain country or region inhabitants. Improvements in economic development is expected to increase host country attractiveness for capitals. Our set of variables includes first host country 's per capita income, *GDPCAP* and growth (*GDPCAPGWTH*). These two variables are expected to increase the net flow of sovereign wealth funds capitals via the so-called wealth effect described in Knill et al. (2012) and Megginson et al. (2013). Following Blonigen et al. (2007), we also include information on skill endowments, (SKILL), measured as the average years of schooling for those over age 25 (Barro and Lee, 2001). The variable is only available at a low frequency with observations reported every five years for 1965-2010. For the sake of consistency with the other series, we interpolate linearly the value for the missing years. Domestic investment made by local SWF is added to the model via the variable *DOMFLOWS*, to test whether cross-border SWFs investments are complementary or substitute to domestic investments. In order to depict the level of economic investment in the host country, we add the variable *TOTINV*, which is the ratio of the total net investment over GDP (see Table 2). Finally, we include two measures of quality of the economic structure of the host country in our model. The first one, *RULELAW*, captures the perceptions of the extend to which agents have confidence in and abided by the rules of the society, and particularly in the quality of contract enforcement and property rights. The second one, *POLSTAB*, assesses the political stability of the country. Those two variables are expected to contribute positively to the country's attractiveness as they depict the quality of the institutional environment for the creditors.

The second set of explanatory variables is related to the financial development of the host country. There are several reasons why better financial development should be beneficial for international investors such as greater credibility, lower transaction costs or higher liquidity to say a few. Three variables are used in the model to capture this dimension of the host economy. First, we consider an indicator measuring the development of the financial intermediary services (*FININT*), equal to the sum of the liquid liabilities and the domestic credit to private sector over GDP.[9] We expect a positive impact of *FININT* on SWFs net flows since a sound financial system could make it easier for an economy to absorb capitals from abroad. As second variable, we consider the MSCI growth (*MSCIGWTH*), which can be adequately summarized as a desire to get into the growing financial market or alternatively to adopt a counter cyclical investment strategy in order to have the opportunity to buy stocks at a lower price ("cheap market effect"). We final financial development variable considered in our regression is the stock market volatility (*STKMKTVOL*), a measure of uncertainty in financial markets.[10] As such, we expect a negative impact of the host country volatility on SWFs net flows. Finally, we include in our regression time-effects to account for the presence of macroeconomic shocks. Descriptive statistics of explanatory variables are displayed in Table 3.

As often done in macroeconomic studies, we lag all explanatory variables of one

[9] We interpol the liquid liabilities variable for 6 countries Bahrain(2004 and 2007), Canada (2009), Chile (2004 to 2009) and Norway (2007 to 2009) and we interpol the domestic credit to private sector variable for 2 countries: Canada (2009) and Norway (2007 to 2009).

[10] We interpol the stock market volatility for 4 countries: Bahrain (2004) and Morocco (2005). .

period to mitigate possible endogeneity issues. Table 2 provides more details about the construction of explanatory variables along with the expected effects, while Table 4 displays the correlation matrix between explanatory variables.

3.3 Spatial interaction matrices

Accounting for spatial interactions in SWFs net flows requires the set up of an interaction scheme, modeled through the use of an interaction (spatial weights) matrix W_n. In this contribution, we use three types of matrices to account for cross-sectional interactions between host nations.

The first considered way of modeling interactions between host countries consists in using geographical neighbors. The associated matrix of interactions is based on the exponential square inverse distance between host countries. To this end, we use the database from the CIA factbook website. In the same spirit of Megginson et al. (2013); Knill et al. (2012); Avendano (2010) and Chhaochharia and Laeven (2008) who analyze the effect of distance between SWF and host nations, we contribute to the existing literature by investigating the effect of distance between host countries themselves on capitals they attract from SWFs. Formally, this geographical interaction matrix is constructed as follows:

$$w_{ij} = \begin{cases} 0 & \text{if} \quad i = j \\ e^{-2d_{ij}} & \text{otherwise} \end{cases} \tag{8}$$

As presented in the introduction, interactions between host countries can be justified by the economic closeness between these countries. Therefore, the second interaction scheme used is based on bilateral trade flows. The elements of this matrix are constructed from bilateral monthly imports and exports of host countries from 1998 to 2003 (72 observations), to avoid any endogeneity issues. These bilateral data come from the IMF database and are expressed in USD. More precisely, the interaction between each pair of host countries is constructed as follows:

$$w_{ij} = \begin{cases} 0 & \text{if} \quad i = j \\ \frac{\overline{X}_{ij}+\overline{M}_{ij}}{\overline{X}_{i.}+\overline{M}_{i.}} & \text{otherwise} \end{cases} \tag{9}$$

where $\overline{X}_{ij} = \frac{1}{T}\sum_{t=1}^{T} X_{ij,t}$ is the average exports over the period from country i to j , $\overline{M}_{ij} = \frac{1}{T}\sum_{t=1}^{T} M_{ij,t}$ represents the average imports of country i from j, $\overline{X}_{i.} = \frac{1}{T}\sum_{t=1}^{T} X_{i.,t}$ are the average exports from country i to the rest of the world while $\overline{M}_{i.} = \frac{1}{T}\sum_{t=1}^{T} M_{i.,t}$ are the average imports of country i from the rest of the world.

Another argument relative to the high level of concentration of investments in international asset allocation is related to the existence of privileged relationship between home country and host country. The literature on foreign direct investment and international asset allocation emphasize the critical role of symmetric information (cost) and familiarity in explaining the spatial allocation of international investments (Beugelsdijk and Frijns, 2010). Accordingly, several studies focusing on behavioral arguments for the existence of the home bias as well as over-investment in specific foreign regions have suggested using language proximity as behavioral proxies (Solnik, 2008; Knill et al., 2012; Chhaochharia and Laeven, 2008). Therefore, the third considered interaction matrix is based on this linguistic proximity. The matrix is constructed using the Common Language index (CL) developed by Melitz and Toubal (2014). Their CL index is based on strictly on exogenous linguistic factors and is constructed from three different indices: Common Official Language (COL), Common Native Language (CNL) and Language Proximity (LP). Their index ranges between 0 and 1. Formally, the elements of this interaction matrix take on the following form:

$$w_{ij} = \begin{cases} 0 & \text{if} \quad i = j \\ CL_{ij} & \text{otherwise} \end{cases} \tag{10}$$

where CL_{ij} is the Common Language index value between host countries i and j.

All interaction matrices are normalized by their spectral radius (see Kelejian and Prucha, 2010).[11]

4 Empirical part

In this section, we identify the factors improving host countries' attractiveness for SWFs' investments as well as the specific role of spatial interactions between host countries regarding their allocation.

Table 5 contains the results of the estimation of model 7. Each column contains the results for an interaction matrix. In the first column, we use the geographic distance interaction matrix; in the second column, results are for the economic proximity based interaction matrix while the third column contains the results when interactions are modeled through linguistic proximity. The Wald est on the Mundlak correction variables (Table 5) shows that these control variables are jointly statistically significant, implying a significant correlation between individual effects and regressors.

Estimation results show a negative and significant spatial autoregressive parameter λ, no matter the interaction matrix used. Economically such a result finds its justification in the 'spatial competition' between countries. Turning to the performance of the different specifications, it is also noticeable that the likelihood value is higher for the linguistic proximity matrix. For the sake of simplicity, in the rest of the paper we will therefore interpret the results obtained with this interaction matrix. We also observe that the estimated coefficient of the IHS transformation, $\hat{\theta}$ is significantly different from zero. This means that the transformed dependent variable is more adequate than the original dependent variable to model the data (see MacKinnon and Magee, 1990). For comparison purposes, estimation results using the original dependent variable, without any transformation are presented in Table 6. We first observe a non-significant spatial autoregressive parameter, no matter the interaction matrix used. Besides, we loose the significance of all but the stock market volatility and the skill variable. We also note that the coefficient associated to the skill variable switched of sign. Finally, we remark that the constant term becomes significant.

[11]Interaction matrices are normalized to ensure the comparability of spatial autoregressive parameters coming from different models.

Interpretation of the results in spatial model is not trivial. Model 2 is an implicit form model estimated by maximum likelihood. However, in order to compute the impact of a change of a regressor on the dependent variable, one needs to rely on the reduced form of the model, which is shown in equation 11.

$$\omega(Y)_{nt} = (I_n - \lambda_0 W_n)^{-1}\left[\sum_{k=1}^{K} X_{nt,k}\beta_{0,k}\right] + (I_n - \lambda_0 W_n)^{-1}U_{nt} \tag{11}$$

According to Pence (2006), two types of impacts can be computed from the reduced form of a model with IHS-transformed dependent variable. Firstly, one can calculate the impacts of a level change of the originally measured dependent variable due to a change in an explanatory variable. In the context of spatial autoregressive models, the matrix of marginal impacts for the k^{th} determinant is

$$\frac{\partial Y_{nt}}{\partial X'_{nt,k}} = \frac{\partial Y_{nt}}{\partial\omega(Y)'_{nt}}\frac{\partial\omega(Y)_{nt}}{\partial X'_{nt,k}} = \frac{\partial Y_{nt}}{\partial\omega(Y)'_{nt}}\left[(I_n - \lambda_0 W_n)^{-1} I_n \beta_{0,k}\right]$$

Referring to Pence (2006), Y_{nt} is the equivalent to the hyperbolic sine of $\omega(Y)_{nt}$. Hence, for individual i at period t, we have:

$$y_{it} = \frac{1}{2\theta}\left[e^{\theta_0\omega(y)_{it}} - e^{-\theta_0\omega(y)_{it}},\right]$$

which implies:

$$\frac{\partial y_{it}}{\partial\omega(y)_{it}} = \frac{1}{2}\left[e^{\theta_0\omega(y)_{it}} + e^{-\theta_0\omega(y)_{it}},\right] \tag{12}$$

This derivative will differ for different values of y_{it}, a property common to nonlinear models. In equation 13, we construct $\partial Y_{nt}/\partial\omega(Y)'_{nt}$, a diagonal matrix where its i^{th} diagonal element corresponds to $\partial y_{it}/\partial\omega(y)_{it}$. To ease the interpretation, we will assess the partial derivative $\partial Y_{nt}/\partial\omega(Y)_{nt}$ at a given value of y_{it}, for instance its median. For this particular case, equation 13 takes the expression shown in equation 14.

$$\frac{\partial Y_{nt}}{\partial\omega(Y_{nt})'} = \frac{1}{2}\begin{pmatrix} \left[e^{\theta_0\omega(y)_{1t}} + e^{-\theta_0\omega(y)_{1t}}\right] & \cdots & 0 \\ 0 & \ddots & \vdots \\ 0 & \cdots & \left[e^{\theta_0\omega(y)_{nt}} + e^{-\theta_0\omega(y)_{nt}}\right] \end{pmatrix} \tag{13}$$

$$\frac{\partial Y_{nt}}{\partial\omega(Y_{nt})'} = \frac{1}{2}\left(e^{\theta_0\omega(y_{med})} + e^{-\theta_0\omega(y_{med})}\right) I_n \tag{14}$$

where $\omega(y_{med})$ is the IHS-transformed dependent variable evaluated at its median value. Finally, we obtain the matrix of impacts (evaluated at the median value of the untransformed dependent variable y_{nt}) which measures the effect on the dependent variable (in levels) of a change in the k^{th} regressor as:

$$\frac{\partial Y_{nt}}{\partial X'_{nt,k}} = S_{n,k}(W_n) = (I_n - \lambda_0 W_n)^{-1}\frac{1}{2}\left[I_n\beta_{k,0}\left(e^{\theta_0\omega(y_{med})} + e^{-\theta_0\omega(y_{med})}\right)\right] \tag{15}$$

As usual for spatial autoregressive models, this impact matrix is full due to the $(I_n - \lambda_0 W_n)^{-1}$ term. The diagonal elements of $S_{n,k}(W_n)$ represent direct effects, i.e. $\partial y_{it}/\partial X_{it,k}$. They are all different from each other due to own-spillover effects, inherently heterogeneous due to differentiated terms in W_n. This heterogeneity of direct effects is what Debarsy and Ertur (2010) call interactive heterogeneity. Off-diagonal terms of $S_{n,k}(W_n)$ represent indirect effects, $\partial y_{it}/\partial X_{jt,k}$, i.e. the effect on the dependent variable of observation i due to a change in the k^{th} regressor of observation j.

By contrast to non-spatial models where the impact of a change of a regressor on the dependent variable is a scalar, in spatial autoregressive models we need to interpret a full matrix. To ease interpretations, LeSage and Pace (2009) propose scalar measures which summarize the information contained in impact matrices. They define the average direct effect as $n^{-1}\operatorname{tr}(S_{nk}(W_n))$, the average total indirect effect as $n^{-1}\iota_n' \left(S_{n,k}(W_n) - \operatorname{Diag}(S_{n,k}(W_n))\right)\iota_n$ where $\operatorname{Diag}(A)$ is a diagonal matrix containing the diagonal elements of A, and the average total impact as $n^{-1}\iota_n' \left(S_{n,k}(W_n)\right)\iota_n$.

The average total impact measures the effect on the dependent variable of an observation of a change of the same amount of the k^{th} regressor in all observations, or alternatively the total effect on all observations of a change in the k^{th} regressor in one observation.

The average total indirect effect measures the impact on all observations but i of a change in the k^{th} regressor in observation i, or alternatively, the impact on an given observation of a change in the k^{th} regressor of the same magnitude in all observations but the concerned one.

The second type of impacts computation suggested by Pence (2006) consists in calculating the effect of a unit change in the k^{th} regressor on the percentage change of the dependent variable. This interpretation is akin to that in a specification where the dependent variable is log-transformed. Remembering the link between the IHS and log transformation, this interpretation is valid for large values of the dependent variable. The matrix of partial semi-elasticities for a SAR specification is presented in equation 16.

$$\frac{\partial \omega(Y_{nt})}{\partial X'_{nt,k}} = (I_n - \lambda_0 W_n)^{-1}\beta_k \approx \frac{1}{\theta_0}\frac{\partial \ln(Y_{nt})}{\partial X'_{nt,k}} \tag{16}$$

For large values of the dependent variable, the partial derivative of the IHS-transformed dependent variable wrt. the k^{th} regressor approximates the partial derivative of the log-transformed dependent variable wrt. the same regressor, ignoring a multiplicative factor.

Properties of this impact matrix in terms of direct, indirect and total effects is the same as for equation 15 and the summary scalar measures of LeSage and Pace (2009) can be used.

To draw inference regarding the statistical significance of the scalar summary measures, one needs their distribution. In this contribution, we follow LeSage and Pace (2009) and construct this distribution using a large number of simulated parameters drawn from the multivariate normal distribution of the parameters implied by the maximum likelihood estimates.

To analyze the effects of the determinants on SWFs net flows, we need to rely on the marginal impacts or semi-elasticities given by expression 15 and 16 respectively and reported in Tables 7 and 8 for the linguistic proximity interaction matrix[12]. We will only interpret the results in terms of semi-elasticities, reported in Table 8 and will give some examples of interpretation with the amounts.

Concerning direct impacts, we provide evidence that on average, SWFs net flows increase with *DOMFLOWS*, *GDPCAP*, *POLSTAB* and *FININT* and decrease with *SKILL* and *STKMKTVOL*.

The positive direct effect of *DOMFLOWS* implies that domestic investments by local SWFs are not made at the expense of foreign investments but act more as a complement. As discussed in the introduction, there has been in some countries a long lasting fear about sovereign wealth funds investments (see, among others Chhaochharia and Laeven, 2008; Karolyi and Liao, 2010; Knill et al., 2012). One way to interpret this result therefore is that political and industrial leaders are more inclined to receive foreign SWF capitals as a local fund actively invests in parallel domestically. The findings regarding to *GDPCAP* and *FININT* show that more mature economies - i.e. economies with higher GDP per capita and sound financial market - tend to attract more investments. Political stability of the host country is also as expected a factor that positively contributes to the attractiveness. Conversely, increased uncertainty in the financial environment as measured by high volatility in the local stock market dampens SWF net flows. Eventually, everything been equal, countries with low skills workers attract more capitals.

Economically, our results mean for instance that the average direct effect of an increase of *GDPCAP* in a host country of \$10 thousand will be an increase of SWFs net flows of on average 4.6%.

We now turn to the indirect impacts. These impacts measure the response of SWF net flows in a given country of a variation (of the same magnitude) of an explanatory variable in all other countries of the sample. Due to negative interactions between host countries, their sign will be opposite to those obtained for direct effects. We note that the average indirect effects associated to *DOMFLOWS*, *GDPCAP*, *POLSTAB*, *SKILL*, and *FININT* are statistically significant.

These results imply that one should expect significant change in the net inflows of investments in a given country if one of the aforementioned variable varies in others countries.

Alternatively, we could also say that impacting one of the 5 variables mentioned above in a single host country will affect the net investments in all other countries. For instance, increasing the GDP per capita of 10 thousand dollars in a country will lead to a cumulated decrease of net investments of 0.76% on average in all other countries.

The presence of significant indirect effects indicate that policy makers should consider the global effect of a policy and not the effect on the host country only. Indeed, shocking a variable in a given country will lead to changes in the attraction of SWFs investments of all the other considered countries. Contrasting with other variables, it worth noting that even though *STKMKTVOL* has a significant negative direct

[12] Marginal impacts and semi-elasticities for the economic proximity interaction matrix are available upon request to the authors but are quantitatively and quantitatively similar.

effect on the attraction of SWFs capital, this variable does not have any significant indirect effect.

Finally, we observe significant total effects for *DOMFLOWS*, *SKILL*, *POLSTAB* *GDPCAP*, *FININT* and *STKMKTVOL*. The sign of these total effects are the same as those of direct effects, but semi-elasticities are of lower magnitude, due to the dampening effect of indirect impacts. For instance, results in Table 8 indicate that if the GDP per capita of a country increases by 10 thousand dollars, the total effect on the SWFs net flows in all countries (including the country where the variation occurred) is a increase of 3.9 % on average.

Marginal impact results, presented in Table 7 are qualitatively equivalent to those obtained using the semi-elasticities interpretations but since these impacts are computed over the original dependent variable (amounts), they depend on the value of SWFs investments, since the model is nonlinear. In this contribution, we decided to analyze the effects at the median value of the investments, namely $125.64 millions.

Concerning direct impact, we provide evidence that on average, SWFs net flows increase with *DOMFLOWS*, *GDPCAP* *POLSTAB* and *FININT* and decrease with *SKILL* and *STKMKTVOL*.

Economically, these marginal impacts mean for instance that the average direct effect of an increase of *GDPCAP* in a host country by ten thousand dollars will increase the SWFs net flows of the considered country by $596.08 millions on average.

As for the semi-elasticities interpretations, we observe significant average indirect effects for *DOMFLOWS*, *GDPCAP*, *POLSTAB*, *SKILL*, and *FININT* variables.

Regarding total effects, we note, as for semi-elasticities, that they are of the same sign as direct effects but with lower magnitude, due to the dampening impact of indirect effects.

Finally, in Figure 2 we represent the distribution of the IHS transformation of SWFs net flows amount with the optimal $\widehat{\theta}$ of the linguistic proximity matrix which is equal to 0.042 (see Table 5). We notice that the distribution of this variable looks much more like a normal distribution than the distribution of the original dependent variable, shown in Figure 1.

5 Conclusion

This paper aims to shed light on the question of why some countries are more attractive for SWFs investments than others. To do so, we develop an original framework that quantifies the specific role of spatial dependence in the location of SWFs' investments on one side and on the other side is robust to outliers while dealing with zero and negative values of investments thanks to the Inverse Hyperbolic Sine transformation to the dependent variable. More specifically, the estimation procedure we propose is a random effects spatial autoregressive panel model with an Inverse Hyperbolic Sine transformed dependent variable, where the possible correlation between individual effects and regressors is accounted for through the use of the Mundlak approach. Three interaction schemes are considered to model interactions between host countries. The first one is based on geographic distance. The second one is the economic proximity which is captured by the relative intensity of bilateral trade between countries. The third one reflects linguistic proximity. Several insights emerge from our analysis. Hence, these results suggest that country-level variables can affect sovereign wealth fund net flows and thereby potentially help to contribute to growth. Countries with higher GDP per capita and Domestic flows tend to attract more SWFs capital. SWFs net flows also appear to depend on the level of the financial development captured by the ratio of financial resources to the private sector over GDP. Besides, better political stability and higher level of financial development of host country contribute positively to SWF's investments net flows, while stock market volatility has the opposite effect. Beyond that, at the system wide level, our findings additionally show that these effects should not be considered in isolation. The spatial autoregressive parameter is negative and significant for the three interaction matrices. In accordance, there will be multilateral consequences in changes in countries? characteristics, as investors tend to reallocate their portfolios in response to financial and economic developments in neighbors countries. The results are consistent with arguments that countries with close economic tights or those sharing cultural similarities are set in competition. One way to interpret this result is that once having the necessary information to invest in a bunch of neighboring countries, investors tend to concentrate their investment in the one providing the most interesting perspectives. That being said, this paper by taking the perspective of the host country does not provide a full insight of the mechanisms at play and especially of how each sovereign wealth funds behave individually. This issue and the necessary adjustment of the spatial models to be used in this case are left for future research. Still, we make clear that the analysis of sovereign wealth funds net flows should take into account country characteristics but also the externalities to other countries. From an econometric perspective, the key insight from this paper is that we can adjust standard panel and spatial panel model estimation procedure to apply of the IHS transformation. As discussed throughout the paper, this transformation is very convenient to deal with the presence of outliers as well as null and negative values of the dependent variable. There are obviously many research in macroeconomics and finance such as modeling of foreign direct investment, trade XX that could benefit from this transformation and the associated estimation procedure we propose.

Aknowledgements

Nicolas Debarsy acknowledges financial support of the Region Centre Val de Loire (APR-IA 2014 00094514). Jean-Yves Gnabo acknowledges financial support of the ARC grant (Projet d'Actions de Recherche Concertées) #13/17-055 granted by the Académie universitaire Louvain. We are grateful to Ingmar Prucha, Fei Jin, Hélène Raymond, Sophie Béreau, Oscar Bernal and the participants of the 15th workshop in spatial econometrics and statistics, the 8th Jean Paelinck Seminar, the 7^{th} French Econometrics Conference, the 26 EC^2 conference in Edimburgh.

Table 1: Descriptive Statistics of the dependent variable ($ MM).

	Average	Std.	Min	Max
Argentina	2.21	4.36	-1.13	10.70
Australia	572.99	774.37	-270.75	1905.39
Austria	-542.52	1495.04	-3558.52	271.56
Bahrain	123.63	188.00	-35.29	470.27
Belgium	317.02	225.49	123.57	595.56
Brazil	550.07	650.10	-74.66	1757.37
Canada	1272.35	1413.31	277.52	4090.62
Chile	22.17	74.06	-67.39	152.82
China	1775.50	3400.16	-3093.65	6434.96
Colombia	2.03	4.53	-0.56	11.22
Czech Republic	32.99	135.99	-159.62	258.97
Denmark	411.80	896.63	-878.44	1770.00
Egypt	-3.48	1054.45	-1772.92	1487.32
Finland	381.68	432.02	42.55	1062.37
France	2172.24	2755.57	-1658.63	6293.18
Germany	4208.06	4144.35	388.17	11182.65
Greece	117.71	124.42	-8.07	327.54
Hong Kong	734.85	1328.26	52.13	3422.90
Hungary	44.47	147.00	-114.71	303.92
Indonesia	-303.27	1324.47	-2977.83	579.24
Ireland	64.64	201.14	-153.66	285.79
Israel	58.58	107.17	-26.19	266.05
Italy	701.52	607.87	141.73	1482.09
Jordan	-224.90	1697.16	-3532.56	1154.33
Malaysia	627.67	768.54	-693.08	1532.99
Mexico	169.89	518.59	-715.90	602.33
Morocco	2.79	6.90	-0.72	16.84
Netherlands	352.82	552.07	-340.57	1286.46
New Zealand	10.43	23.41	-28.17	34.47
Norway	18.95	32.67	-3.23	82.65
Peru	4.05	12.42	-5.37	28.87
Philippines	16.14	22.51	-2.54	55.27
Poland	51.99	134.19	-50.37	319.01
Portugal	185.66	257.39	-98.61	551.87
Russia	8150.54	18098.23	-999.35	44940.69
Singapore	79.72	101.22	-51.79	186.42
Spain	788.20	1134.91	-502.92	2509.58
Sweden	441.16	401.93	-302.48	927.05
Switzerland	3760.68	3205.73	1260.00	9891.49
Thailand	345.78	671.99	-513.81	1468.03
Turkey	124.52	210.30	-24.85	500.61
United Kingdom	6549.32	6699.62	471.87	15066.14
United States	15625.00	13656.56	6089.76	37272.06

Table 2: Description of the explanatory variables

Notation	Definition	Sources	Expected Sign
	Economic Development and Structure		
GDPCAP	GDP per capita of the host country is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant (thousand) 2005 U.S. dollars.	WBI	(+)
GDPCAPGWTH	Annual percentage growth rate of GDP per capita of the targeted country based on constant local currency. Aggregates are based on constant 2005 U.S. dollars. GDP per capita is gross domestic product divided by midyear population. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources.	WBI	(+)
SKILL	Average years of schooling for those over age 25 of the targeted country (Barro and Lee, 2001).	Barro and Lee (2001)	(+/-)
DOMFLOWS	Domestic SWFs net flows amount.	FactSet	(+/-)
TOTINVT	Total value of the gross fixed capital formation and changes in inventories and acquisitions less disposals of valuables for a unit or sector divided by the GDP.	IMF	(+)
RULELAW	Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.	WGI	(+)
POLSTAB	Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism.	WGI	(+)
	Financial Development		
FININT	Sum of the liquid liabilities and the domestic credit to private sector over GDP (financial resources provided to the private sector) of the targeted country (Demirguc-Kunt and Levine, 1996).	GFDI	(+)
MSCIGWTH	National MSCI index growth rate of the targeted country. The variable is multiplied by 100 for scaling adjustment.	MB	(-)
STKMKTVOL	Average of the 360-day volatility of the national stock market index of the targeted country.	GFDI	(-)

Notes: WBI=World Bank Indicator database; GFDI=Global Financial Development Index database; WGI=World Governance Indicator, Polity IV project database; IMF=International Monetary Fund; MB = Macrobond.

Table 3: Descriptive Statistics for explanatory variables

	Average	Std.	Min	Max
DOMFLOWS ($ bil.)	0.21	1.32	-1.68	14.14
GDPCAP ($ M.)	22.56	17.96	1.12	67.80
GDPCAPGWTH (%)	3.10	2.64	-4.58	13.57
POLSTAB (*100)	21.83	93.44	-239.00	166.00
RULELAW (*100)	80.88	94.01	-95.00	200.00
SKILL (*100)	958.85	227.23	357.60	1330.46
TOTINVT (%)	23.28	4.79	14.26	43.78
FININT (%)	173.08	87.36	37.57	434.77
MSCIGWTH (%)	3.51	11.50	-36.78	25.20
STKMKTVOL	19.52	7.40	7.56	44.96

Table 4: Correlation matrix of explanatory variables

		1	2	3	4	5	6	7	8	9	10
1	MSCIGWTH	1									
2	TOTINVT	-0.12	1								
3	STKMKTVOL	-0.21	-0.12	1							
4	POLSTAB	-0.09	-0.03	-0.29	1						
5	DOMFLOWS	0.06	0.19	-0.04	-0.03	1					
6	RULELAW	-0.10	-0.11	-0.36	0.84	-0.05	1				
7	GDPCAP	-0.14	-0.15	-0.23	0.72	-0.06	0.86	1			
8	GDPCAPGWTH	0.31	0.24	0.18	-0.32	0.15	-0.48	-0.49	1		
9	SKILL	-0.15	-0.21	-0.14	0.55	0.00	0.64	0.65	-0.21	1	
10	FININT	-0.17	0.18	-0.32	0.49	0.11	0.65	0.55	-0.22	0.35	1

Table 5: SWFs net flows amount, IHS transformation used

	Geo. Distance	Economic Prox.	Language Prox.
	Economic Development		
CST	-32.033	-10.633	-13.483
	(44.295)	(47.758)	(46.563)
DOMFLOWS	6.400**	6.334**	6.644**
	(2.892)	(2.953)	(2.963)
GDPCAP	10.277**	10.063**	11.079**
	(4.413)	(4.509)	(4.422)
GDPCAPGWTH	-0.568	-0.504	-0.546
	(2.568)	(2.621)	(2.625)
POLSTAB	0.372*	0.421*	0.421*
	(0.220)	(0.225)	(0.225)
RULELAW	-0.243	-0.359	-0.264
	(0.536)	(0.547)	(0.547)
SKILL	-0.455**	-0.475**	-0.515***
	(0.182)	(0.186)	(0.189)
TOTINVT	0.527	0.754	0.687
	(1.891)	(1.930)	(1.932)
	Financial Development		
FININT	0.539**	0.544**	0.540**
	(0.229)	(0.234)	(0.235)
MSCIGWTH	-0.625	-0.645	-0.636
	(0.678)	(0.692)	(0.693)
STKMKTVOL	-1.424*	-1.429*	-1.447*
	(0.810)	(0.829)	(0.833)
	Spatial autocorrelation		
λ	-0.474*	-0.255*	-0.351**
	(0.246)	(0.143)	(0.166)
	IHS transformation		
θ	0.043***	0.042***	0.042***
	(0.005)	(0.005)	(0.005)
Obs	258	258	258
Wald statistic	41.984	41.640	44.322
Wald Pval	0.000	0.000	0.000
llike	-1968.837	-1969.033	-1968.490
Time fixed effects	yes	yes	yes

Notes: ***, ** and * indicate significance level at the 1%, 5% and 10% level. Standard errors reported between brackets.

Table 6: SWFs net flows amount, original dependent variable

	Geo. Distance	Economic Prox.	Language Prox.
	Economic Development		
CST	-5979.734*	-5728.517*	-5915.666*
	(-3311.664)	(-3191.836)	(-3336.546)
DOMFLOWS	-274.697	-280.763	-273.101
	(459.179)	(469.874)	(452.604)
GDPCAP	66.881	47.063	97.060
	(-80.328)	(-63.442)	(-121.712)
GDPCAPGWTH	-150.489	-134.270	-144.296
	(-398.135)	(-288.929)	(-507.785)
POLSTAB	45.011	42.927	45.187
	(-33.155)	(-30.982)	(-33.503)
RULELAW	-22.502	-22.484	-22.689
	(-106.6779)	(-151.538)	(-74.426)
SKILL	5.681*	6.996	4.290*
	(3.307)	(5.280)	(2.547)
	Financial Development		
FININT	27.317	21.066	26.791
	(23.819)	(16.914)	(23.512)
MSCIGWTH	86.226*	85.806*	84.719*
	(-47.601)	(-47.779)	(-47.788)
STKMKTVOL	6.221***	12.144***	6.347***
	(2.125)	(4.356)	(2.159)
TOTINVT	-242.066	-240.255	-236.986
	(-2218.060)	(-1130.270)	(-2130.231)
λ	-0.269	-0.146	-0.154
	(0.192)	(0.106)	(0.111)
Obs	258	258	258
llike	-2491.317	-2490.138	-2491.306
Time fixed effects	yes	yes	yes

Notes: ***, ** and * indicate significance level at the 1%, 5% and 10% level. Standard errors reported between brackets.

Table 7: Average Impacts on SWFs net flows amount ($ Millions), evaluated at median (125.64 $ Millions)

	Language Prox.		
	Direct Impact	Indirect Impact	Total Impact
	Economic Development		
DOMFLOWS	35.746**	-5.795**	29.952**
	(8.007;69.618)	(-13.439;-0.615)	(6.905;57.989)
GDPCAP	59.608***	-9.663**	49.945***
	(18.424;110.463)	(-21.292;-1.1)	(15.873;93.019)
GDPCAPGWTH	-2.939	0.477	-2.463
	(-25.499;20.842)	(-3.746;4.081)	(-21.279;17.351)
POLSTAB	2.267*	-0.368*	1.9*
	(0.281;4.48)	(-0.818;-0.01)	(0.236;3.778)
RULELAW	-1.421	0.23	-1.191
	(-6.553;3.348)	(-0.594;1.196)	(-5.635;2.79)
SKILL	-2.771***	0.449**	-2.322***
	(-5.27;-0.717)	(0.06;0.955)	(-4.511;-0.609)
TOTINVT	3.695	-0.599	3.096
	(-14.984;19.874)	(-3.431;2.296)	(-12.359;16.794)
	Financial Development		
FININT	2.905**	-0.471**	2.434**
	(0.859;5.399)	(-1.058;-0.058)	(0.712;4.506)
MSCIGWTH	-3.424	0.555	-2.869
	(-10.133;2.647)	(-0.418;1.823)	(-8.38;2.221)
STKMKTVOL	-7.783*	1.262	-6.521*
	(-16.582;-0.149)	(-0.03;3.095)	(-14.066;-0.124)

***, ** and * indicate significance level at the 1%, 5% and 10% level. 10% confidence interval reported between brackets.

Table 8: Semi-Elasticities (%)

	Language Prox.		
	Direct Impact	Indirect Impact	Total Impact
		Economic Development	
DOMFLOWS	0.28**	-0.045**	0.234**
	(0.061;0.547)	(-0.105;-0.005)	(0.053;0.454)
GDPCAP	0.466***	-0.076**	0.391***
	(0.141;0.868)	(-0.167;-0.009)	(0.117;0.733)
GDPCAPGWTH	-0.023	0.004	-0.019
	(-0.201;0.164)	(-0.029;0.032)	(-0.167;0.135)
POLSTAB	0.018*	-0.003*	0.015*
	(0.002;0.035)	(-0.006;0)	(0.002;0.03)
RULELAW	-0.011	0.002	-0.009
	(-0.051;0.026)	(-0.005;0.009)	(-0.044;0.022)
SKILL	-0.022***	0.004**	-0.018***
	(-0.041;-0.005)	(0;0.007)	(-0.036;-0.005)
TOTINVT	0.029	-0.005	0.024
	(-0.118;0.154)	(-0.027;0.018)	(-0.097;0.131)
		Financial Development	
FININT	0.023**	-0.004**	0.019**
	(0.006;0.042)	(-0.008;0)	(0.005;0.035)
MSCIGWTH	-0.027	0.004	-0.022
	(-0.079;0.021)	(-0.003;0.014)	(-0.066;0.017)
STKMKTVOL	-0.061*	0.01	-0.051*
	(-0.13;-0.001)	(0;0.024)	(-0.11;-0.001)

Notes: ***, ** and * indicate significance level at the 1%, 5% and 10% level. 10% confidence interval reported between brackets.

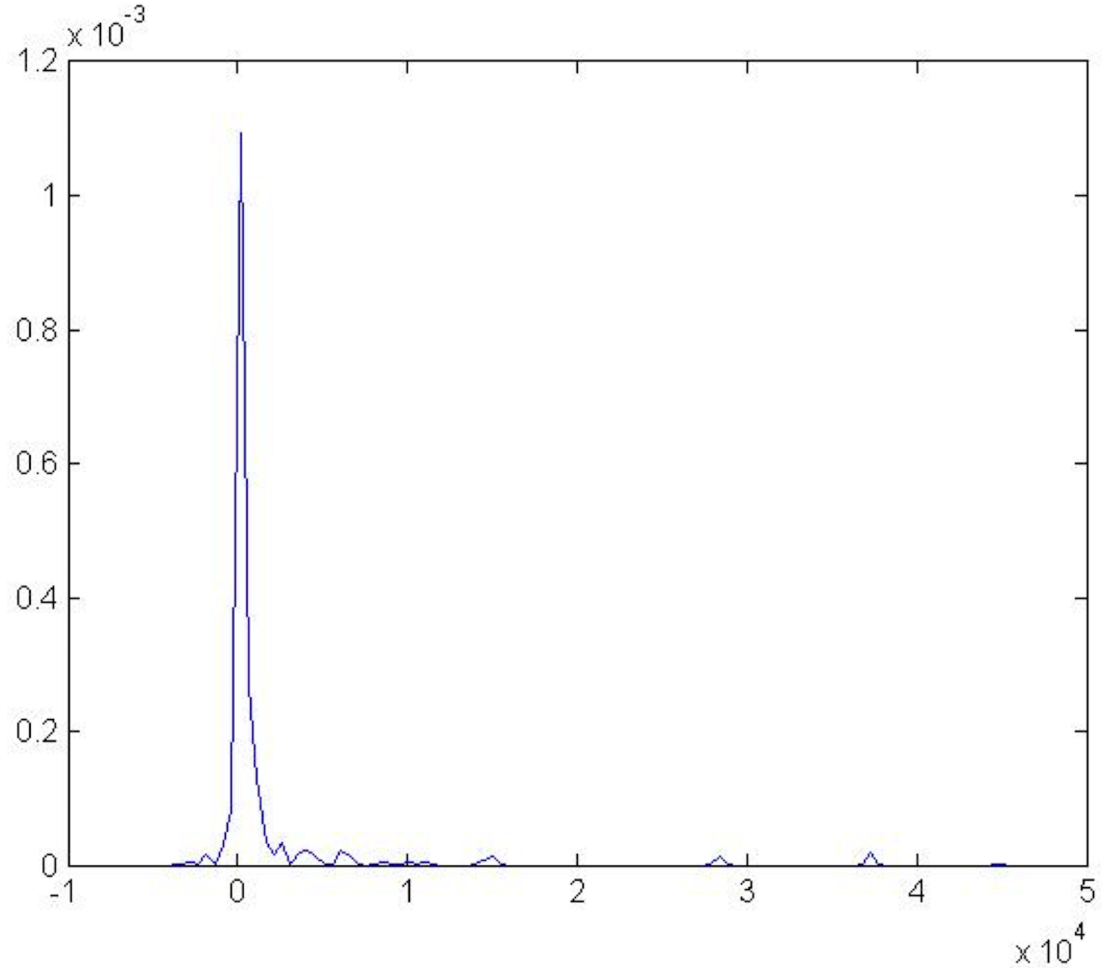

Figure 1: Foreign SWFs net flows amount

References

Avendano R. 2010. Sovereign wealth fund investments: From firm-level preferences to natural endowments. *Working paper, Paris School of Economics* .

Aviat A, Coeurdacier N. 2007. The geography of trade in goods and asset holdings. *Journal of International Economics* **71**: 22–51.

Barro R, Lee JW. 2001. International data on educational attainment updates and implications. *Economic Papers* **53**: 541–563.

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

Bernstein S, Lerner J, , Schoar A. 2009. The investment strategies of sovereign wealth funds. *NBER Working paper* .

Beugelsdijk S, Frijns B. 2010. A cultural explanation of the foreign bias in international asset allocation. *Journal of Banking & Finance* **34**: 2121–2131.

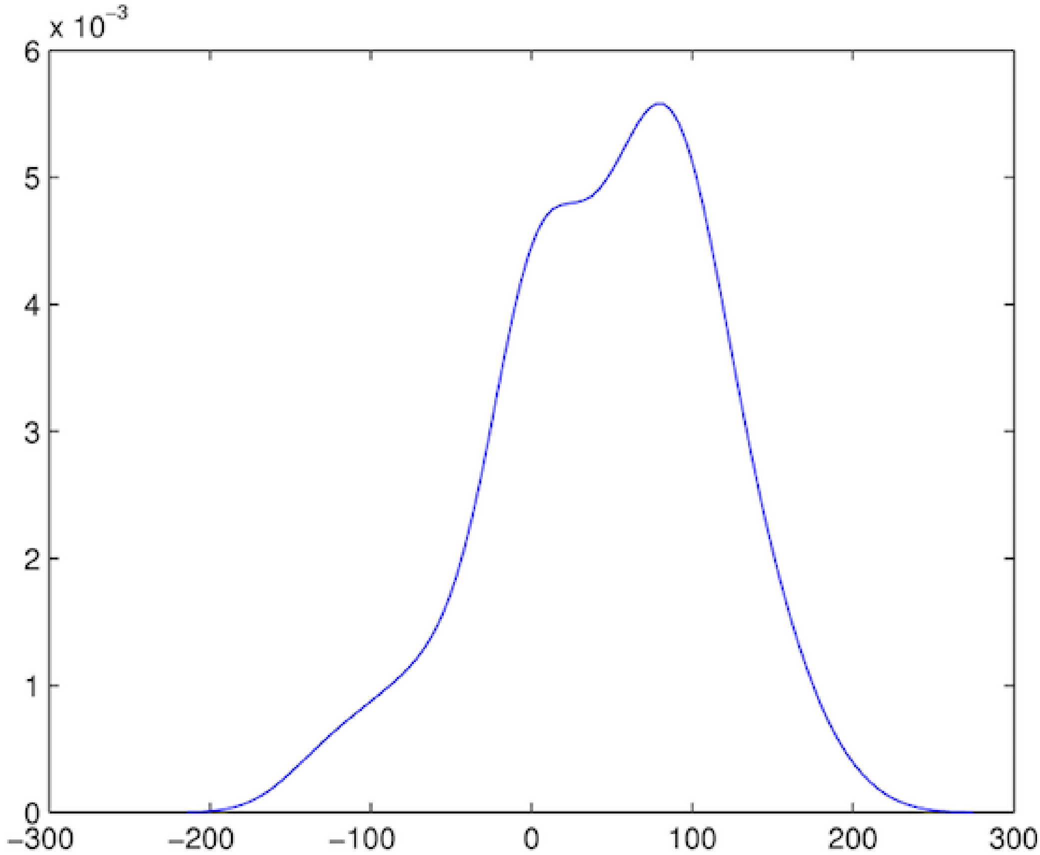

Figure 2: Foreign SWFs net flows amount IHS transformed with $\hat{\theta} = 0.0420$

Bickel P, Doksum K. 1981. An analysis of transformations revisited. *Journal of the American Statistical Association* **76**: 296–311.

Blonigen B, Davies R, Waddell G, Naughton H. 2007. Fdi in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review* **51**: 1303–1325.

Bortolotti B, Fotak V, Megginson WL. 2013. The sovereign wealth fund discount: evidence from public equity investments. *Baffi Center Research Paper* .

Breusch TS. 1987. Maximum likelihood estimation of random effects models. *Journal of Econometrics* **36**: 383–389.

Burbidge JB, Magee L, Robb AL. 1988. Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association* **83**: 123–127.

Carroll C, Dynan KE, Krane S. 2003. Unemployment risk and precautionary wealth: evidence from households' balance sheets. *Review of Economics and Statistics* **85**: 586–604.

Chhaochharia V, Laeven L. 2008. Sovereign wealth funds: Their investment strategies and performance. Discussion Paper 6959, C.E.P.R.

Debarsy N. 2012. The mundlak approach in the spatial durbin panel data model. *Spatial Economic Analysis* **7**: 109–131.

Debarsy N, Ertur C. 2010. Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics* **40**: 453–470.

Demirguc-Kunt A, Levine R. 1996. Stock market development and financial intermediaries: Stylized facts. *World Bank Economic Review* **10**: 291–321.

Dewenter KL, Han X, Malatesta P. 2010. Firm values and sovereign wealth fund investments. *Journal of Financial Economics* **98**: 256–278.

Disdier AC, Head K. 2008. The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and statistics* **90**: 37–48.

Eaton J, Tamura A. 1994. Bilateralism and regionalism in japanese and us trade and direct foreign investment patterns. *Journal of the Japanese and international economies* **8**: 478–510.

Fernandes N. 2011. Sovereign wealth funds: Investment choices and implications around the world. *Working paper* .

Fotak V, Bortolotti B, Megginson W, Miracky W. 2008. The financial impact of sovereign wealth fund investments in listed companies. *Working paper* .

Frankel JA, Stein E, Wei SJ. 1997. *Regional trading blocs in the world economic system.* Peterson Institute.

Hau H. 2001. Location matters: An examination of trading profits. *Journal of Finance* **56**: 1959–1983.

Hesse H, Sun T. 2009. Sovereign wealth funds and financial stability-an event study analysis. *IMF Working paper* **9**: 239.

John J, Draper NR. 1980. An alternative family of transformations. *Journal of the Royal Statistical Society, Serie C* **29**: 190–197.

Johnson NL. 1949. Systems of frequency curves generated by methods of translation. *Biometrika* : 149–176.

Kapteyn A, Panis C. 2003. The size and composition of wealth holdings in the united states, italy, and the netherlands. *NBER Working paper 10182* .

Karolyi G, Liao RC. 2010. What is different about government-controlled acquirers in cross-border acquisitions? *Working paper University of Cornell* .

Kelejian HH, Prucha IR. 2010. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics* **157**: 53–67.

Kennickell AB, Sunden AE. 1997. Pensions, social security, and the distribution of wealth. *Federal Reserve Board of Governors. Finance and Economics Discussion Series* **55**: 1–41.

Knill A, Lee BS, Mauck N. 2012. Bilateral political relations and sovereign wealth fund investment. *Journal of Corporate Finance* **18**: 108–123.

Kotter J, Lel U. 2011. Friends or foes? target selection decisions of sovereign wealth funds and their consequences. *Journal of Financial Economics* **101**: 360–381.

Kristjánsdóttir H. 2005. Determinants of foreign direct investment in iceland. *Working paper* .

Lee LF, Yu J. 2010a. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics* **154**: 165–185.

Lee LF, Yu J. 2010b. Some recent developments in spatial panel data models. *Regional Science and Urban Economics* **40**: 255–271.

Lee LF, Yu J. 2012. Spatial panel: random components versus fixed effects. *International Economic Review* **53**: 1369–1412.

LeSage J, Pace RK. 2009. *Introduction to spatial econometrics*. CRC Press Taylor and Francis Group.

Lin X, Lee LF. 2010. GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *Journal of Econometrics* **157**: 34–52.

Linders GJ, De Groot HL. 2006. Estimation of the gravity equation in the presence of zero flows .

MacKinnon J, Magee L. 1990. Transforming the dependent variable in regression models. *International Economic Review* **31**: 315–39.

Megginson WL, Fotak V. 2014. Rise of the fiduciary state: A survey of sovereign wealth fund research. *Available at SSRN 2432623* .

Megginson WL, You M, Han L. 2013. Determinants of sovereign wealth fund cross-border investments. *The Financial Review* **48**: 539–572.

Melitz J, Toubal F. 2014. Native language, spoken language, translation and trade. *Journal of International Economics* **93**: 351–363.

Mundlak Y. 1978. On the pooling of time series and cross section data. *Econometrica* **46**: 69–85.

Pence KM. 2006. The role of wealth transformations: An application to estimating the effect of tax incentives on saving. *Contributions in Economic Analysis & Policy* **5**: 1–24.

Portes R, Rey H. 2005. The determinants of cross-border equity flows. *Journal of international Economics* **65**: 269–296.

Raballand G. 2003. Determinants of the negative impact of being landlocked on trade: an empirical investigation through the central asian case. *Comparative Economic Studies* **45**: 520–536.

Raymond H. 2010. Sovereign wealth funds as domestic investors of last resort during crises. *International Economics* **123**: 121–160.

Reynolds A, Shonkwiler J. 1991. Testing and correcting for distributional misspecifications in the tobit model: An application of the information matrix test. *Empirical Economics* **16**: 313–23.

Solnik B. 2008. Equity home bias and regret: An international equilibrium model. *Available at SSRN 828405* .

Wansbeek T, Kapteyn A. 1982. A class of decompositions of the variance-covariance matrix of a generalized error components model. *Econometrica: Journal of the Econometric Society* : 713–724.

Yen ST, Jones A. 1997. Household consumption of cheese: An inverse hyperbolic sine double-hurdle model with dependent errors. *American Journal of Agricultural Economics* **79**: 246–251.

Yu J, Walsh JP. 2010. Determinants of foreign direct investment; a sectoral and institutional approach. *IMF Working paper* **10**: 87.

Zhang J. 2014. *Foreign direct investment, governance, and the environment in China: Regional dimensions.* Palgrave Macmillan.