A pure measure of home bias*

Ian A Cooper[†] Piet Sercu[‡] Rosanne Vanpée[§]

August 30, 2016

Abstract

The literature on international equity holdings distinguishes between home bias (overweighting of home stocks) and foreign bias (relative underweighting for more 'distant' countries). The two biases can be integrated into one distance-based model. We define pure home bias as the excess of home bias relative to this model, and find pure home bias only in emerging markets. Countries with high tax rates and low credit standing have higher pure home bias, and more development comes with lower distance aversion. Methodologically, the choice of portfolio bias measure matters. We find the best measure to be a covariance-based measure relative to the world average.

JEL classification: G15, G18, G30, G38, F3

Keywords: Home bias, foreign bias, pure home bias, distance, deadweight costs, economic and financial development, reputation,

*This is a revised version of a paper previously called: Is equity home bias different from foreign bias? The authors thank Prof. Anil Mishra (University of Western Sydney) and Andrés Fernndez Martin (Inter-American Development Bank) for kindly sharing their data. We are also grateful to Ingrid Werner, Andrew Karolyi, Paolo Maio, Chandra Thapa, Kris Boudt, Martina Vandebroek, Tom Vinaimont and Hans Degryse, as well as participants in seminars at the University of Strathclyde, Hong Kong City University, the Central University of Finance and Economics (Beijing), the 2015 Multinational Finance Society in Cyprus, the Brussels 3L workshop and the 2016 American Finance Association in SF, for useful comments and discussions. We accept full responsibility for any remaining errors.

[†]London Business School, London NW1 4SA, England, UK; +44-207-000-7000; icooper@london.edu.

[‡]KU Leuven, Faculty of Economics and Business, Naamsestraat 69, 3000 Leuven, Belgium; +32-16-326756; piet.sercu@kuleuven.be.

[§]KU Leuven, Faculty of Economics and Business, Warmoesberg 26, 1000 Brussels, Belgium; +32-02-6081432; Rosanne.vanpee@kuleuven.be .

1 Introduction

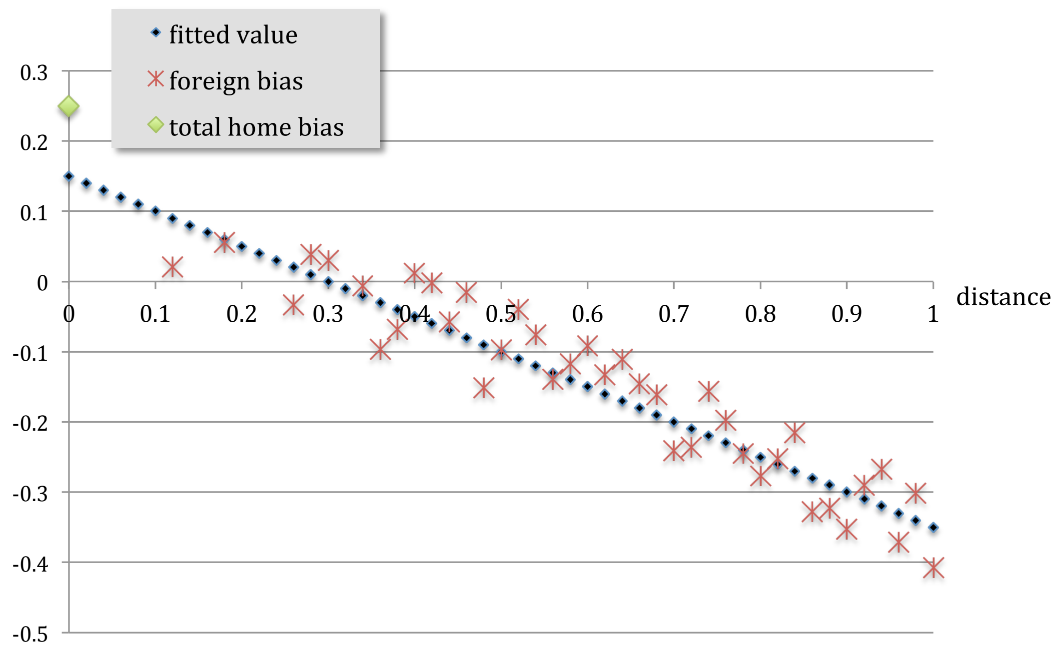
The literature on international equity holdings distinguishes between home bias and foreign bias. Home bias refers to the observation that investors overweight their home-country stocks so much that the forgone marginal benefit to increased international diversification looks far greater than any observable costs of holding foreign equities, such as withholding taxes. Foreign bias, in contrast, refers to the observation that investors assign higher portfolio weights to foreign countries that are 'closer' or more similar to their home country, even though these are the countries where the diversification benefits tend to be smaller. 'Closeness' here refers to a combination of geographic proximity, intensity of high information and trade flows, historical or cultural links, and other common traits like similarities in income levels.

The phenomena of home and foreign bias have typically been treated separately (*e.g* Chan, Covrig and Ng (2005), Lau, Ng and Zhang (2010), Bekaert and Wang (2009)). In this study we integrate them, on the basis of the observation that in both biases the choices depend on the 'closeness' between the investing country and the country of investment: the home country is the closest country and also exhibits the highest bias. Thus the fact that foreign bias is related to closeness might in itself already explain much of, and perhaps all of, the home bias. In short, our key question is to what extent the home country differs from a hypothetical foreign country that is just extremely close. If so, the new question is whether that difference is ubiquitous or, instead, mostly arises in certain countries.

In this paper, the research question is expressed in terms of distance rather than closeness. For our purposes, the former is the negative of the latter, and both are defined very broadly. So distance covers not just physical distance and items related to unfamiliarity, but also dissimilarities in any variables that seem to affect portfolio decisions. That is, in our terminology, items like divergences between taxes may make the countries more 'distant' in the very wide sense of being less close or less similar – provided the variable which experiences this difference matters empirically in portfolio choice. It is more crucial that all potential ingredients are differences vis-a-vis the home country, so that by construction the home base itself has distance zero.

The basic question whether 'home' is more than a zero-distance country leads to two con-

Figure 1: Home bias explained by distance versus pure home bias



cepts. One is 'pure' home bias, which is the amount of home bias in excess of the level that would be predicted for a foreign country with zero distance, as obtained from an observed relation between foreign bias and an index of distance. The latter relation produces the second parameter that is of interest to us, namely the country's distance aversion.

Distance aversion and pure home bias are distinct concepts, but both refer to a regression of portfolio bias on distance. Figure 1 illustrates the idea for the case where bias is measured as the gap between a destination country's weights in an investor's chosen portfolio versus in the world market portfolio. For a representative investor domiciled in country i there is a relation between the distance of various foreign countries, j, and the investor's excess holdings of stocks in those countries. The points marked by asterisks in the graph refer to various foreign countries, a few of which here are overweighted in the investor's portfolio while most are underweighted. In this example the portfolio bias weights are plotted against distance, our X-coordinate, defined so that the home country has zero distance. The asterisks – the foreign-bias observations, *not* including the home bias – lie on a noisy line, which is downward sloping, which indicates that more distant countries have lower (and usually more negative) excess weights. The slope coefficient measures the country's distance aversion.

We repeat that the home country itself is not used in this foreign-bias regression. Distance and distance aversion, in our view, sum up foreign bias, but we are of course agnostic whether also home bias is solely driven by distance. For this reason we should not let home bias affect our estimate of distance aversion – nor of the composition of the distance index, for that matter.

Returning to the example, we see that the line cuts the Y-axis at 0.15, which is the amount of bias predicted for a zero-distance foreign country based on the country's foreign bias pattern. But the home country itself has a distance of zero as its X-coordinate. In the example, the total home bias overweight is 0.25, shown as the diamond on the Y-axis. Of this 0.25 overweight, 0.15 is explained by the absence of distance. We conclude that in this example the domestic investments still exceed what can be traced to the absence of distance, and we measure the pure home bias as the residual part, $0.25 - 0.15 = 0.10$.

To implement our idea, we estimate, for each country, the foreign bias relation that can be inferred from its foreign portfolio positions. We measure distance as a weighted average of various differences vis-a-vis home, with weights common to all source and destination countries, but we let each origin country have its own sensitivity to the synthetic distance index ('distance aversion'). The intercept of this foreign-bias regression has the interpretation of the portfolio bias for a putative foreign country with a 'distance' of zero. In the example we assume no non-linearities, but in the real data we of course test for that. We then evaluate whether the home investments differ from the intercept (that is, whether pure home bias is different from zero). Lastly, we investigate what country characteristics are related to pure home bias.

In terms of empirical results, our contributions are related to four themes: (i) choosing between alternative measures of portfolio bias, (ii) assessing distance aversion and pure home bias, a process which in turn requires the identification of a common synthetic measure of distance, and (iii)–(iv) identifying the determinants of distance aversion and pure home bias, respectively. The remainder of this introduction provides more details about each of these.

Regarding theme (i), we consider four measures of portfolio bias which have been used in prior studies. Two turn out to be flawed in terms of empirical validity. One of these, unexpectedly, is the popular gap between the weights country i assigns to the various destination countries' assets and the destination countries' weights in the world capitalisation. This measure, we find, is too similar across investing countries: it is mostly driven by the destination country's relative market size (the biggest entry in the difference, common to all home countries) and is hardly affected by the investing country's choices, a small number. So it tells much about the destination countries, but little about differences in the behaviour of different investors towards those destination countries. The Sercu and Vanpée (2008) covariance-gap measure likewise mostly reflects the destination country's variance rather than the investing country's own choices.

A third measure, the difference of the logs of observed and benchmark weights, raises serious conceptual doubts and produces empirical results that make little sense. Our preferred measure compares the covariances of a destination country's return with the investing country's portfolio and the world market portfolio ('covar-W'). The first of these covariances, multiplied by risk aversion, translates into a set of corresponding net[1] excess returns as expected by the investing country for each asset; and the second covariance then reflects a world average implied expectation. Home bias, as measured in this way, is both statistically and economically significant.

As an ingredient in our estimation of distance aversion and pure home bias, theme (ii) in our list of contributions, we identify an index of distance that explains covar-W well. The components in the synthetic distance index (*i.e.* the measure of distance on the horizontal axis of Figure 1) include taxes, capital import restrictions abroad, geographical distance, (minus) trade intensity, and relative market size. Some measures identified in earlier papers, such as language and common currency, do not correlate meaningfully with the covar-W measure of foreign bias; currency does affect the traditional measure, though. Using covar-W as the measure of bias and its associated measure of distance, we then document how for almost two-thirds of our 41 sample countries, we can accept the Null of no pure bias: the level of home bias is not clearly higher than if the home country were a foreign country with zero distance. This group consists of developed markets. The U.S. has the lowest pure home bias, presumably

[1]Net means after any objective cost, shadow cost, or behavioural bias that may apply.

because its population of home stocks is unusually large and diverse. This ties in with Lévy (2013)'s diagnosis that, for the U.S., home bias is not an economically serious problem. Our numbers however show that one cannot generalise Lévy's conclusion to all countries. Judging by the point estimates, even midsize EU countries have sizeable implied pure expectations gaps, and those for Egypt, Turkey or Russia are extremely large.

Regarding themes (iii) and (iv), we examine the determinants of distance aversion, and find that distance aversion is lower for more developed markets. When we similarly explore home-country characteristics that go with a country's pure home bias, we find it is associated with two seemingly contradicting groups of variables. A low degree of development and sophistication in an origin country increases pure home bias, but so do a high sovereign spread and high taxes at home – variables that should discourage home investments rather than encourage them. So it looks as if these characteristics cause home bias via *foreign* reluctance to invest in the home market, not so much via domestic portfolio choice. One should therefore regard a low total home bias in country i not just as the result of the residents' interest in foreign assets ('push' items, mostly reflected in foreign bias) but also of the foreigners' interest in i's home assets ('pull' characteristics, mostly affecting pure home bias).

We proceed with a brief discussion of related earlier work, followed by a methodology section, the foreign-bias regressions, and the analysis of country characteristics associated with pure bias and distance aversion. The final section sums up.

1 Literature review

The extensive literature on home bias and foreign bias is summarized in, *e.g* Cooper, Sercu, and Vanpée (2013) and many other papers cited therein. In this section we accordingly discuss only that part of the literature which relates directly to the present paper.

Our work is closely related to distance-based or gravity models of foreign bias, in particular Portes and Rey (2005). Portes and Rey estimate a version of the foreign-bias regression based on portfolio flows. They regress the foreign equity portfolio flows, in dollars, from country i to country j in period t on the size of country i's equity market, the size of country j's equity market, the geographical distance between the countries, measures of the information transfer between the countries, measures of the efficiency of transaction technology, and cyclical variables. They find that these 'gravity' variables explain a large proportion of the variation in portfolio flows, and that portfolio diversification measures are not significant.[2]

Our work is also related to Bekaert and Wang (2009), who investigate both home bias and foreign bias. To that end they construct a measure of foreign bias stripped of its interdependency with home bias. Specifically, they adjust the foreign bias measure for the amount of equity invested in foreign markets, which reflects the degree of home bias. In this study we do almost the opposite: we strip the amount predicted by the foreign bias relation from the home bias measure. Thus our focus is on pure home bias rather than on improved measurement of foreign bias, which is the focus of Bekaert and Wang.

Chan, Covrig and Ng (2005), followed by Lau, Ng and Zhang (2010), also distinguish between a domestic bias and a foreign bias and they identify the determinants of these biases. Based on a sample of equity holdings of mutual fund portfolios of 26 countries they show that stock market development and familiarity variables explain the investment bias, but the impact on domestic and foreign bias is different. If a destination country is more remote from the rest of world, domestic investors hold more of that country's stocks, while foreign investors invest less in that country. When a destination country has a more developed financial market, foreign investors invest more in that country's stocks, while domestic investors invest less. Chan *et al.* define the domestic and foreign bias as the log-ratio of the actual weight of a country's stocks in the mutual fund portfolio and the weight of this country's stock in the world market portfolio.

Both Bekaert–Wang and Chan *et al.* run separate home bias and foreign-bias regressions. Their results show important differences. For example, a tax variable is highly significant in

---

[2] Although the work of Portes and Rey provides the primary motivation for our distance-based analysis of the foreign bias, there is one aspect of their study that should be interpreted with caution when compared with our results. They report very high R^2s, but those cannot be compared with the R^2s we report in this study. Their regressions have unadjusted and unscaled dollar flows as dependent variables and use the size of the home and foreign markets as dependent variables. The portfolio bias measures we use in the foreign-bias regression, in contrast, are already scaled by the size of the home market (because they are portfolio weights or transformations thereof) and benchmarked on the basis of the size of the foreign market (because they are relative to a portfolio benchmark) – numbers that appear as explanatory variables in the Portes and Rey regression. Therefore, we expect lower R^2s, simply because the effects of the dollar scale of the home and foreign markets are removed from our measures.

Bekaert and Wang, but not in Chan *et al.* Also, different variables appear to be important for foreign bias than for home bias: for instance, measures of return are significant in the foreign-bias regressions but not in the home-bias regressions. Importantly, in both studies, the measures of correlation are highly significant in the foreign-bias regressions but not in the home bias ones. This suggests that diversification considerations play a part in the foreign bias. However, the way diversification is included is largely *ad hoc*, namely as a single correlation and with an additive role. Overall, the differences in the results show that the empirical determinants of home and foreign biases depend on the portfolio bias measure used and possibly also the data period.

This analysis is related to the literature which finds that distance appears to dominate diversification benefits in forming international equity portfolios (Tesar and Werner (1995)). Bekaert and Wang similarly find that the diversification effect goes the wrong way in the foreign-bias regressions: countries with higher correlations have lower foreign bias. Our results suggest that including covariances as part of the left-hand side variable is useful in understanding the structure of international equity holdings. This could be because investors do take the various levels of covariance into account, or alternatively it may be because the measures based on just portfolio holdings suffer from a potential scaling problem, as discussed in Section 2.4.1. However, because countries with low correlations are also more 'distant' countries, one has to be careful in disentangling the effects of correlation from those of distance. A covariance-based variable does exactly that.

A final strand of the literature which is related to our work is the attempt to identify the benefits investors gain by holding local stocks. If underdiversification is rational, the foregone diversification from not holding 'distant' stocks should be offset by a gain from holding local stocks. It could be, for example, that investors have more information about local stocks and this results in a higher expected return. However, attempts to measure the gain resulting from such information have found little systematic evidence for such outperformance by local investors (*e.g* Seasholes and Zhu (2010)).

The literature related to the specific explanatory variables we use is discussed in Section 2.2, along with the data. First, in the next section, we outline the model and the measures of portfolio bias.

2 Test set-up

We first propose our general way to address the research question, and then turn to the variables that act as right- and left-hand side covariates.

2.1 General estimation procedure

In this section we describe our distance-based empirical model of the foreign bias for country i. There are N countries. Suppose that K distance variables may affect the foreign holdings chosen by investors located in country i, and denote the kth such measure of difference between country i and foreign country j by X_{ijk}. By construction, $X_{iik} = 0$. We are less concerned about whether this difference reflects true costs, such as incremental tax burdens, or simply captures behavioural biases with no underlying economic costs, as long as the synthetic measure of distance picks up covariates that empirically do matter.

A weighted average of all of these distance measures is used as a scalar index that sums up the distance between country i and foreign country j. Let w_k denote the weights in the index (constrained to be equal across countries), and $\overline{X}_{ij}$ the scalar index level for the pair of countries (i, j):

$$\overline{X}_{ij} := \sum_{k=1}^{K} w_k X_{ijk}, \quad \{j = 1, \cdots, N\}, \text{ with } \sum_{k=1}^{K} w_k = 1. \tag{1}$$

Lastly, let Y_{ij} be some measure of portfolio bias in the holdings by investors i of assets from country j.[3] We assume, initially, a linear model for the foreign bias of country:

$$Y_{ij} = a_i + b_i \overline{X}_{ij} + e_{ij}, \quad i \neq j, \tag{2}$$

where a_i is a constant that is specific to country i. The parameter b_i describes the 'distance aversion' or slope of the foreign bias of country i.

In the above, b_i is written as country specific. We examine whether the slopes could be the same across countries, but that Null is rejected. In all regressions we impose common

[3] The four measures we consider are discussed in Section 2.4.

values on the weights w_k assigned to the various sub-measures X_{ijk} within $\overline{X}_{ij}$. Subject to that constraint, we want to find the weighting scheme that best explains the observed bias. So our procedure is to solve, jointly for all home countries i, the following least-squares problem, non-linear in the estimands:

$$\text{minimise}_{\forall a_i,\, b_i,\, w_k} \sum_i \sum_{j \neq i} \left[Y_{ij} - a_i - b_i \cdot \sum_k w_k X_{ijk} \right]^2 \text{ subject to } \sum_k w_k = 1. \tag{3}$$

Recall that the home country bias measures are omitted from this regression.

Since the pure home bias measure relies on an estimated zero-distance fitted value (the intercept), the assumption of linearity is crucial. To assess the adequacy of that assumption, we compare Equation (2) to several variants, adding either distance squared as a regressor, or its inverse, or both, or adopting an exponential version of the original equation. To evaluate the Null that pure home bias is zero (that is, all home bias is explained by the absence of distance), we test whether each intercept as estimated above is statistically different from the total home bias $Y_{ii}^{(.)}$.[4] In addition, we test what home-related variables Z_i can explain the observed pure home bias. The Null hypothesis is that the home bias equals the population intercept, so that any estimated gap (pure home bias) deviates from zero only because both the distance aversion and the weights within the distance measure are measured with noise. So in the regression

$$PHB_i = A + d\, Z_i + u_i, \text{ where } PHB_i := Y_{ii}^{(.)} - \hat{a}_i, \tag{4}$$

the Null of no pure home bias predicts $A = 0$ and $d = 0$.[5]

---

[4]For the exponential, the appropriate measure is $Y_{ii}^{(.)} - e^{\hat{a}_i}$.

[5]A potential problem with the two-step regression procedure is that the second stage dependent variable, $\hat{a}_i$, includes a measurement error from the estimation of the foreign bias whose dispersion is not explicitly taken into account when setting the standard errors for the step-two estimates. But one should bear in mind that the two-step procedure still delivers 41 intercepts from as many regressions, not one set of parameters all generated by a single regression. These 41 intercepts should, therefore, still contain a decent sample of estimation errors which are taken into consideration in stage two. In addition, the two-step approach offers a clean separation between the estimation of pure home bias and its explanation. Lastly, one-step regressions are possible only if the home-bias equation is subtracted from the foreign-bias equation:

$$Y_{ij}^* := Y_{ij} - Y_{ii}, \tag{5}$$

so that the revised home-bias measure for $j = i$ equals zero, as in covar-L. But this leads back to the validity trap we just encountered: the home bias (the Y_{ii} term) dominates, so that the new left-hand variables become too similar across investing countries. For these reasons we just apply the two-step approach.

Because we find that distance aversion b_i is country specific, we also investigate what country characteristics are associated with the degree of distance aversion. This is done via a regression:

$$\hat{b}_i = B + fZ_i + v_i. \tag{6}$$

We now proceed with a discussion of the candidate components of the overall distance index (Section 2.2), their expected role in the overall measure (Section 2.3) and the candidate measures of portfolio bias (Section 2.4).

2.2 Component measures of distance

We consider 11 potential entrants X_{ijk} into the synthetic distance function, and construct them so that a higher X_{ij} means that assets from j become more attractive to investors from i. These variables can be classified as explicit costs and barriers to capital inflows, familiarity indicators, and proxies for governance and financial market sophistication. Many such variables have been used in studies of home and foreign bias. For example, Bekaert and Wang and Chan *et al.* use a total of 34 variables between them. These fall into 6 major groups: development, information/distance, capital controls/openness, tax, investor protection/governance, and diversification. Within each group, many of the variables are highly correlated, limiting the usefulness of employing multiple proxies for the same basic characteristic. Collinearity produces unstable weights within the synthetic distance index that often defy economic logic. This could still be shrugged off as unimportant; but more crucially for our purpose, imprecise slopes also generate imprecise intercepts,[6] which then makes it too easy to accept the Null of zero home bias.

For these reasons, we adopt a stepwise selection procedure, as detailed in Section 3.3.1. The most important measures, in prior studies, turn out to be development, distance, common

---

[6] In $\mathbf{Y} = \alpha\mathbf{e} + \mathbf{XB} + \mathbf{u}$, with $\mathbf{e}$ denoting a vector with elements all equal to unity, the OLS variance of the estimated intercepts equals

$$\text{var}(\widehat{\alpha}|X) = \frac{1}{n}\left(1 + \overline{X}'\widehat{\text{var}}(X)^{-1}\overline{X}\right)\text{var}(\varepsilon),$$

with $\widehat{\text{var}}(X) = \sum_i (X_i - \overline{X})(X_i - \overline{X})'/n$. Hence, α will be estimated imprecisely when the regressors are highly collinear. We use robust standard errors rather than OLS ones, but the same logic applies: imprecise slope estimates generate imprecise intercepts.

Table 1: **Regressors: definition, source, and descriptive statistics**

Variables related to foreign bias (standardized)		source	Period	Mean	Std dev	Min	Max
Tax-div$_{ij}$	Withholding tax rate on dividends (host – home)	PriceWaterhouseCoopers' Worldwide	2001–2012	0.01	0.15	– 0.34	0.47
Tax-capgain$_{ij}$	Withholding taxes on capital gains (host – home)	Anil Mishra	2001–2011	–0.01	0.20	–0.45	0.45
Capin$_{ij}$	Controls on incoming capital	Schindler (2009) and Fernandez, Rebucci and Uribe (2014)	2001–2012	0.20	0.27	0.00	0.92
Dist$_{ij}$	Log of geographical distance between home and host	http://www.cepii.fr/	constant	3.74	0.44	2.24	4.30
D-trade$_{ij}$	Trade separateness between home and host	calculated on the basis of IMF Direction of Trade Statistics (trade), IMF Financial Statistics and World Bank (GDP, current account)	2001–2012	0.90	0.09	0.34	1.00
Language$_{ij}$	Common language dummy		constant			0.00	1.00
Currency$_{ij}$	Common currency dummy	IMF, Annual Report on Exchange Arrangements and Exchange Restrictions	2001–2012			0.00	1.00
Flist$_{ij}$	Foreign listings from host on home market	Sarkissian and Schill (1998, 2003, 2014) for foreign listings, WFE for domestic	1998/2003/2006	0.98	0.05	0.48	1.00
Size$_{ij}$	log-ratio of GDP of home and host country	World Bank	2001–2012	0.01	0.69	–2.10	2.10
Soph$_{ij}$	Financial market sophistication index (host – home)	Financial Development Reports by the World Economic Forum	2001–2012	0.02	1.34	–3.28	3.28
Gov$_{ij}$	Government effectiveness (host – home)	World Bank Governance Indicators	2001–2012	0.03	1.23	-3.19	3.19

Variables related to home bias (home-country characteristics, not standardised)		source	Period	Mean	Std dev	Min	Max
Tax-div$_i$	Withholding tax rate on dividends	PriceWaterhouseCoopers' Worldwide	2001–2012	0.14	0.08	0.00	0.47
Tax-capgain$_i$	Withholding taxes on capital gains	Anil Mishra	2001–2011	0.15	0.15	0.00	0.45
Capout$_i$	Controls on outgoing capital	Schindler (2009) and Fernandez, Rebucci and Uribe (2014)	2001–2011	0.33	0.37	0.00	1.00
Size$_i$	Log of GDP of the home country	World Bank	2001–2012	11.65	0.51	10.72	13.21
HDI$_i$	Human Development Index	United Nations Development Programme	2001–2012	0.81	0.10	0.48	0.94
GDP/cap$_i$	Log of GDP per capita	World Bank	2001–2012	4.17	0.50	2.67	5.00
Gov$_i$	Government effectiveness (home)	World Bank Governance Indicators	2001–2012	0.99	0.87	–1.19	2.43
Remote$_i$	Log of average distance from home to host countries	http://www.cepii.fr/	constant	3.70	0.21	3.44	4.09
Patriot$_i$	Indicator of patriotism in home country	World Values Survey, How proud are you to be [Nationality]?	1999–2014	3.42	0.30	2.62	3.91
Forbes$_i$	Weight in Forbes Global 2000 List over world market weight m_i	Forbes Global 2000	2012	0.80	0.41	0.00	1.75
Spread$_i$	Interest rate spread over 10-y U.S. Treasury	World Bank staff calculations (Datastream), CDS spreads	2007–2012	3.01	12.22	–0.25	148.69
Rating$_i^2$	Moody's country rating (Aaa=24 to C=4), squared	www.moodys.com	2001–2012	397.30	174.23	16.00	576.00

Key The table lists, in the first panel, the variables $X_{..k}$ used to construct the index $\overline{X}_{ij} = \sum_{k=1}^{K} X_{ijk}$ that sums up the difference ('distance') between countries i and j. In the second panel we list the variables $Z_{i,k}$ used to explain distance aversion and pure home bias.

language, capital controls, tax, and governance.[7] The variables we use, listed in Table 1 with source and period data and some descriptive statistics, cover these categories. They are discussed at greater length in Appendix A.

2.3 Expected signs for 'push' and 'pull' variables

For variables that are natural measures of distance, we expect a negative coefficient in the foreign-bias regression; and as the synthetic measure of distance $\overline{X}_{ij}$ is preceded by a negative b_j, that expectation translates into a positive weight inside $\overline{X}_{ij}$. For other variables, such as size and governance, which are differences between home and foreign country characteristics, we define all measures so that a higher value corresponds to destination country j being less attractive to origin country i's residents. Yet we place no *a priori* restriction on the sign of its weight, for the following reason. A country like the United States is larger than all other countries. On these grounds, one expects the U.S. to be a preferred destination for all other countries and to have an unusually large weight, all else being equal, in the foreign portfolios. However, size makes the U.S. attractive also to a its own home investors – an opposite force that may push out the foreigners. It is hard to foretell which consideration will dominate, so it is hard to formulate *a priori* expectations about the sign of the coefficient on the difference in size in the foreign-bias regression. A similar issue arises with any variable where $X_{ijk} = -X_{jik}$, like the difference in the quality of governance. In contrast, when a natural measure of distance is used like geographical distance or trade intensity, the distance between i and j is also the distance between j and i. So if a country like the U.S. is close (and therefore attractive) to some other foreign country, then that second country is also close and attractive to the U.S. They both like each other's assets, instead of both going for just one of the two. As a result, there are no opposite forces, so no problem of combining 'push' and 'pull' effects arises.

For the measures where there is no clear prediction what the weight will be, the estimate's sign tells us which side's view turns out to dominate: the country of origin ('push') or of destination ('pull'). A positive weight corresponds to a situation where the home country's

[7] Diversification has also often been found to be important, but our covariance measure of portfolio bias captures that directly.

aversion to investing in, *e.g* smaller or poor-governance countries trumps the destination country's preference for going for larger or well-run countries, and *vice versa*.

2.4 Candidate measures of home bias

Our work requires data on both international portfolio holdings and country index returns. We use actual portfolio weights based on the international portfolio holdings data from the Coordinated Portfolio Investment Survey (CPIS) provided by the IMF. Because the CPIS has been conducted on an annual basis since 2001, we have annual portfolio weights for the period 2001–2012. Out of the 75 countries participating to the CPIS in 2012, only 42 could be retained in our sample due to data missing in Datastream or shortcomings of the CPIS dataset. We partially correct for third-party holdings or round tripping by reallocating the reported investments in financial offshore centres over the sample countries in proportion to the foreign investments of these centres, following Sercu and Vanpée (2008). Monthly equity returns for our 42 sample countries from January 1994 until December 2012 refer to the Morgan Stanley International Country Indices, also retrieved from Datastream. The risk-free rate is the 3-month U.S. Treasury Bill rate and the world market return is proxied by the return of the MSCI World Index.

We now describe and critically discuss the candidate left-hand side variables, the measures of portfolio bias. Each takes the form of a difference between a characteristic of the position held by investor i in assets from country j and a benchmark number for that characteristic.

2.4.1 The traditional weight-gap measure

The most common and traditional measure is based directly on the portfolio holdings themselves, and is defined as the difference between the actual portfolio weight of country j in the portfolio chosen by the investor located in country i (eq_{ij}) and the weight of country j in the world market portfolio (m_j):

$$\text{(Traditional:)} \qquad Y_{ij}^{(1)} := eq_{ij} - m_j. \tag{7}$$

The advantage of the traditional measure is its simplicity. The thorny issue is whether this number can be interpreted properly without scaling, and if not, how to scale it. For example, assume Sweden invests $eq_{jj} = 30\%$ at home whereas the world average investor puts just $m_j = 1\%$ into those same assets. What, then, is the 'comparable' number for the home bias of the U.S. if U.S. stocks have, say, a world weight of 35%? Is the U.S. investor really equally biased if investing 64% at home (so that the weight gap is again 29%)?

While many authors think that is correct, others argue that a 29% gap is huge for Sweden compared to the U.S.: 29%/1% equals 29 while 29%/35% equals a mere 0.82; so their preferred measure would be $(eq_{ij} - m_j)/m_j$ or eq_{ij}/m_j. Others criticise this. Since the maximum ratio for the U.S., in the example, is $100/35 = 2.8$, by the eq/m measure Swedes would be considered even more home biased than a hypothetically 100% biased American even if they invested as little as 3% at home. Clearly, something is wrong here.

Accordingly, some scale the gap by $(1 - m_j)$. In that view, one lets the Swedes start from their 1% world weight, and one sees how much of the remaining 99% is assigned to home assets, and similarly for the U.S. For Sweden, the answer is $29/99 = 29\%$, a relatively low number compared to the U.S., whose excess holdings of home assets use up $29/65 = 44\%$ of the remaining room.

Getting the scaling right is thorny, in short. Yet it is crucial for our purposes. Suppose the true model is $Y/Z = a + bX + \mathrm{e}$; then the regression $Y = c + dX + u$ is a mis-specification for $Y = cZ + dXZ + eZ$, and both the slopes and the intercepts will be badly affected when countries have different scales Z. Allowing for country-specific slopes and intercepts could pick up some of the scaling, but the consequence still is that the intercepts would mix pure home bias with scaling effects.

Regardless of the scaling, an additional issue is whether one can judge a measure of under-diversification without reference to risk. A domestic investment of 90% has a rather different impact when the home country has a large, diversified stock market instead of just a few mining stocks. The usual way of controlling for risk, by adding a world-market correlation as a regressor, would fail to pick this up if, as usual, home bias is pronounced.[8]

---

[8]The riskiness of an asset j in a portfolio p is measured by its covariance with that specific portfolio; so one should not use the world portfolio here, and not scale the covariance into a correlation.

2.4.2 The log-ratio measure

As a second candidate measure, we consider the log-ratio investment bias, as adopted by, *e.g* Chan, Covrig and Ng (2005) and also applied in, among others, Lau *et al.* (2010). This measure is calculated as the natural logarithm of the actual portfolio weight of country j in investor i's portfolio scaled by the benchmark weight of country j:

$$\text{(Log-ratio:)} \qquad Y_{ij}^{(2)} := \log \frac{eq_{ij}}{m_j}. \tag{8}$$

This measure obviously adopts the scaling by m_j, which is not necessarily the best choice, as we just argued: by that logic, Swedes investing 4% at home are deemed to be more biased than Americans holding a fully local portfolio. The log transform adds issues of its own. For home bias, the ratios are often substantially above unity and therefore sensitive to the basis; taking logs does attenuate this. But for foreign bias, which is core in our analysis, virtually all of the ratios are well below unity, so that taking logs sends these numbers to very negative domains. The resulting outlier problem spirals out of control for the many zero entries (*e.g* zero reported Ghanaian investments in Russia, etc.), so that one either loses those observations, or has to Winsorize them to some arbitrary (but still very negative and influential) level.

2.4.3 Covariance measures of bias: covar-L and covar-W

The third measure we consider is the covariance-based measure proposed by Sercu and Vanpée (2008), which takes into account the relative diversification benefits of different foreign countries as well as portfolio holdings. This measure is the difference between the covariances of country j's equity returns, r_j, from the returns on two portfolios, r_{p_i} and r_{p_j}, p_i being the portfolio held by the reference investor i and p_j the portfolio held by investor j (the stock's home or local investor):

$$\text{(Covar-L:)} \qquad Y_{ij}^{(3)} \; := \; \text{cov}(r_j, r_{p_i}) - \text{cov}(r_j, r_{p_j}), \tag{9}$$

In the label we attach to this measure, covar-L, the L stands for local, which refers to the portfolio that is providing the benchmark.

A difference of covariances, according to portfolio theory, should mirror a difference in expectations about net excess returns, after any deadweight costs or behavioural effect that may arise when i or j holds asset j, like taxes and brokerage fees, information costs, shadow costs of binding capital restrictions and biases in the perception of the expectation. If, *e.g* French assets (j) have a lower covariance vis-a-vis the Swedes' (i) portfolio than vis-a-vis France's total portfolio, we conclude that for these assets the Swedes, being outsiders, face higher costs and may also be negatively biased, while French investors face lower costs and may also be biased in favour of their home assets.

Finally, we consider the Cooper–Kaplanis (1986) covariance measure, which chooses as its benchmark the world market portfolio rather than the portfolio of the foreign investor:

$$\text{(Covar-W:)} \qquad Y_{ij}^{(4)} \quad := \quad \text{cov}(r_j, r_{p_i}) - \text{cov}(r_j, r_w). \tag{10}$$

In this expression, the second covariance, the covariance with the world market portfolio return, reflects an average of the various expected net returns, after deadweight costs, across countries.[9]

Covar-L and covar-W, being proportional to an expected-return spread, need no further standardisation to make them comparable between countries. They also take into account covariance risks with other assets, in a way that is consistent with portfolio theory. A 70% unscaled home bias is deemed to be much worse if, for instance, the home asset is very volatile, or otherwise more risky, than in the case the home asset is safe. Also, when bias is measured as $eq_{ij} - m_k$, there is a direct mechanical link between home and foreign bias: by construction, when foreign countries are under-represented, the home country must be over-represented.[10] But for measures based on covariances, there is no such necessary direct link between the two biases. A downside is that, being an estimate, a covariance-gap measure is a constructed variable, calculated from a sample using a model rather than directly observed.

To sum up, on *a priori* grounds, the covariance measures make more sense, while the traditional measure, and especially the log-ratio, seem to raise many issues. In Section 3.2 below, we show that covar-W seems to emerge as the statistically and economically more

[9] This does not assume the World CAPM holds. An asset's world covariance always is a weighted average of the asset's covariances with each of the national portfolios. As such, it reflects a weighted average of the expectations in the various investor countries. The world CAPM would assume these expected net returns to be the same for all investors. We do not need any such assumption.

[10] For that reason, Bekaert and Wang (2009) develop a portfolio holdings measure of foreign bias which includes an adjustment for the degree of home bias.

Table 2: **Correlation matrix for the explanatory variables**

PANEL A: Correlation matrix for the distance variables

	Tax-capg	Tax-div	Capin	Dist	D-trade	Langu	Currncy	Flist	Size	Soph	Gov
Tax-capg	1.00										
Tax-div	-0.19	1.00									
Capin	-0.06	0.19	1.00								
Dist	0.01	0.10	0.26	1.00							
D-trade	0.04	0.06	0.14	**0.54**	1.00						
Langu	-0.01	-0.01	-0.03	0.04	0.23	1.00					
Currncy	-0.01	0.07	0.15	0.36	0.28	0.01	1.00				
Flist	0.05	0.05	0.09	0.26	**0.52**	0.21	0.16	1.00			
Size	-0.31	0.04	0.07	0.01	-0.16	-0.01	0.00	-0.25	1.00		
Soph	-0.05	-0.07	0.30	0.01	-0.04	-0.02	0.00	-0.11	0.24	1.00	
Gov	-0.06	0.02	0.46	0.01	-0.03	-0.02	0.00	-0.08	0.14	**0.84**	1.00

PANEL B: Correlation matrix for the determinants of pure home bias

	Tax-capg	Tax-div	Capout	Size	Remote	HDI	GDP/cap	Soph	Patriot	Forbes	Spread	Rating[2]
Tax-capg	1.00											
Tax-div	0.16	1.00										
Capout	0.11	-0.12	1.00									
Size	0.13	-0.04	-0.07	1.00								
Remote	-0.16	-0.13	0.41	-0.05	1.00							
HDI	-0.06	0.19	**-0.68**	0.22	-0.37	1.00						
GDP/cap	-0.03	0.20	**-0.68**	0.27	-0.40	**0.96**	1.00					
Soph	-0.25	-0.14	-0.36	0.22	-0.12	**0.52**	**0.63**	1.00				
Patriot	0.06	0.10	0.40	-0.28	0.37	-0.45	-0.44	-0.28	1.00			
Forbes	0.19	0.20	-0.44	0.34	**-0.52**	**0.52**	**0.63**	0.40	-0.36	1.00		
Spread	-0.04	0.17	0.02	-0.18	0.03	-0.10	-0.11	-0.46	0.11	0.10	1.00	
Rating[2]	-0.11	0.10	**-0.61**	0.31	-0.40	**0.83**	**0.87**	**0.76**	-0.35	**0.54**	-0.36	1.00

meaningful measure.

3 Empirical results

3.1 Summary statistics for regressors

Averages, standard deviations, and high-low values can be found in Table 1, in Section 2.2 above. Table 2 adds the correlations between the variables. For the foreign-bias regressions (Panel A in Table 2), two groups are highly correlated (trade distance, geographical distance and foreign listings; and governance and sophistication), so within those two groups there will be some pruning.

For the variables that are considered to explain pure home bias (Panel B of Table 2), problems proliferate, with unsurprisingly high correlations among all variables measuring economic and financial development and with the index of the country's score on outward openness to capital. Accordingly, we will define 5 separate regression specifications, including, in turn, each of the highly similar regressors.

3.2 Properties of the portfolio bias measures

In Section 2.4 above, we discussed four measures of bias on *a priori* grounds. Now we turn to empirical issues regarding the validity of these measures. We first look at the quality of every measure as an input in the foreign-bias regression, and then at the quality of the resulting pure home bias estimates.

3.2.1 Selecting a measure of portfolio bias (1): Validity as a regressand

Empirically, two measures suffer from a capital flaw: they produce almost exactly the same measure of portfolio bias for destination country j regardless of the investor country i. This is clearest for the traditional portfolio holdings measure, the difference between the portfolio weights eq_{ij} held by investor i and benchmark weights m_j. For all countries other than the home one, these benchmark weights m_j are much larger, and have much greater variation, than i's portfolio holdings eq_{ij}; and they are the same for every investor investing in a specific foreign country. Most of the variation in this portfolio bias measure is, therefore, independent of the investor i; so $Y_{ij}^{(1)}$ tells a lot about world-market weights m_j, but almost nothing about i's choices eq_{ij}.

The same issue arises with the covar-L measure of bias for country j, $cov(r_j, r_{p_i}) - \text{cov}(r_j, r_{p_j})$. For emerging markets the second term is close to the variance of the local index, and it is large, so the measure is again dominated by the benchmark part rather than by the home countries' choices. Hence, for both the traditional and covar-L measures, the dependent variables are very highly correlated for different investors. The size of the commonality is enormous: the correlation between $Y_{i.}$ and $Y_{k.}$, measured for two different investors i and k and averaged over all such pairs, amounts to 0.977 for the traditional measure and to 0.975 for the covar-L gap. In short, these are not N series with different observations of foreign bias, but N repetitions of the same benchmark vector (market weights m_j or variance var(r_j)), up to a minuscule variation.

For this reason, we discard the traditional portfolio holdings and covar-L measures. In the rest of this paper we proceed with the remaining measures, covar-W and the log-ratio. The latter comes up with a much lower average correlation between all country pairs' bias numbers (0.468, after deleting all observations with zero holdings). Covar-W, lastly, does best at avoiding double counting, with an average correlation of 0.302.

3.2.2 Selecting a measure of portfolio bias (2): Statistical and economic meaningfulness of total and pure home bias

One additional criterion in choosing between portfolio bias measures is the economic and statistical interpretability of the results. Table 3 presents estimates of home bias for the covar-W and log-ratio measures. The countries have been arranged by covar-W-based home bias, and that covariance has been multiplied by risk aversion, set at three,[11] to provide an impression of the corresponding mean–variance-implied expectations gap. The table also provides a preview of the estimates of pure home bias, which are relevant to the choice of a portfolio bias measure.

Looking first at total home bias, we note that the numbers extracted from the two measures are only weakly related, as shown in the leftmost graph in Figure 2. Their correlation is 0.49. We discuss then separately, as either they measure different phenomena or at least one of them is quite noisy.

An attractive feature of the rescaled covariance measure is that the resulting home bias can be interpreted as a gap between the *p.a.* implied expectations about the home country, compared to the world's average expectation about that market. In general, developed countries have expectations-gap home biases that are low, while we see high numbers for younger and smaller financial markets. There is substantial right-skewness, as illustrated by the divergence between the median (9.1%) and the mean (18%), or the difference between Q3–Q2 (18%) versus Q2–Q1 (4.5%). The gap ranges from a mere 30 bp (U.S.) to 75% (Russia). That is, for the U.S., the index's covariance with the U.S. investor's portfolio hardly differs from its covariance with the world portfolio; but for Russia, with its huge home bias, its historic average volatility exceeding 0.5, and its low correlation with the other markets, the covariance gap rises to about 0.25, imputing an expectations gap of 75%.

The figures are not just quite diverse, but also economically significant. Taking, conserva-

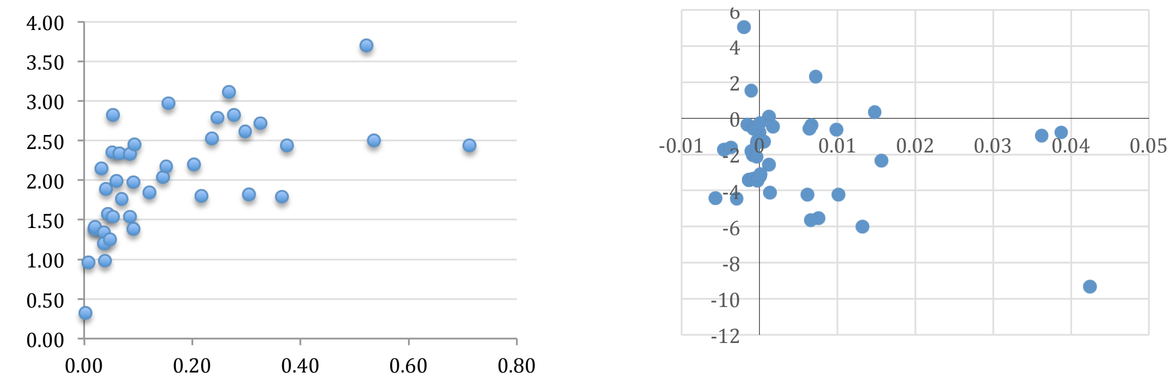
[11] This is about the seminal Friend and Blume (1975) estimate ('well in excess of one and probably in excess of two'). For portfolio volatilities of 0.15 to 0.18 it predicts plausible risk premia of 6.75%–9.75%.

Table 3: **Total home bias and pure home bias for the covar-W and logratio measures**

	covar-W ×3				*logratio*			
	total home bias)		pure home bias		total home bias		pure home bias	
	estim	t-stat	estim	t-stat	estim	t-stat	estim	t-stat
US	0.003***	3.43	0.000	0.03	0.32***	12.92	-0.25	-0.40
UK	0.008***	3.95	-0.005	-0.22	0.96***	94.13	-0.54	-0.81
Switzerland	0.019***	9.44	0.006	0.24	1.38***	43.86	-1.30*	-1.84
Netherlands	0.020***	8.69	-0.008	-0.35	1.41***	34.85	-0.57	-0.99
Denmark	0.032***	9.43	-0.004	-0.17	2.15***	60.38	-1.36	-1.58
Canada	0.036***	12.10	-0.016	-0.67	1.34***	37.77	-0.36	-0.38
France	0.037***	9.16	-0.011	-0.56	1.20***	58.98	-1.82***	-3.04
Japan	0.039***	8.85	0.013	0.68	0.98***	90.79	0.09	0.10
Belgium	0.041***	6.81	0.000	0.00	1.89***	150.10	-0.78	-1.36
Australia	0.044***	9.46	-0.014	-0.70	1.57***	35.88	-3.42***	-3.30
Germany	0.048***	10.95	-0.003	-0.17	1.25***	213.56	-1.25**	-2.60
Portugal	0.052***	11.64	-0.003	-0.10	2.35***	12.60	-3.46***	-2.94
New Zealand	0.053***	11.56	-0.011	-0.47	2.82***	44.73	1.53	0.96
Italy	0.053***	11.57	-0.004	-0.17	1.54***	47.77	-2.14**	-2.49
Norway	0.060***	9.91	-0.008	-0.30	1.99***	12.88	-3.34***	-3.69
Austria	0.065***	10.28	-0.009	-0.38	2.34***	49.52	-2.04***	-2.81
Sweden	0.069***	12.92	0.001	0.06	1.76***	59.39	-3.10***	-3.92
Chile	0.085***	14.11	-0.047*	-1.95	2.33***	26.08	-1.74	-1.00
Spain	0.085***	18.97	0.014	0.60	1.54***	50.45	-4.12***	-3.94
Hong Kong	0.091***	20.07	-0.038	-1.73	1.38***	13.78	-1.62	-1.55
Singapore	0.091***	18.05	-0.059***	-3.33	1.97***	17.75	-4.41***	-4.79
Israel	0.093***	13.14	0.064***	2.65	2.45***	17.29	-4.22	-2.68
South Africa	0.121***	21.63	-0.030	-1.08	1.85***	39.06	-4.46*	-2.05
Finland	0.146***	20.59	0.069***	2.88	2.04***	69.66	-5.65***	-4.77
Mexico	0.152***	19.88	-0.021	-0.76	2.17*	2.13	5.05***	4.47
Czech Rep	0.156***	14.46	0.013	0.47	2.97***	41.30	-2.56**	-2.27
Malaysia	0.203***	17.27	-0.014	-0.61	2.20*	1.90	-3.41***	-3.02
India	0.217***	20.18	0.075***	2.85	1.80***	14.06	2.29	1.37
Greece	0.236***	16.55	0.106***	3.86	2.53**	2.35	-4.24***	-2.94
Colombia	0.246***	29.08	0.069***	2.45	2.79***	14.97	-0.38	-0.23
Hungary	0.268***	15.54	0.001	0.03	3.11***	78.14	-3.22***	-2.92
Egypt	0.277***	20.95	0.163***	5.57	2.82***	26.13	-2.34	-1.32
Poland	0.298***	23.88	0.067***	2.55	2.61***	24.65	-0.58	-0.66
Brazil	0.305***	22.61	0.079***	3.20	1.82***	14.77	-5.54***	-3.32
Argentina	0.326***	24.23	0.138***	5.10	2.72***	31.36	-5.98***	-3.13
Korea	0.366***	25.83	0.154***	6.52	1.79***	16.20	0.34	0.28
Thailand	0.375***	38.41	0.018	0.63	2.44***	18.87	-0.46	-0.39
Venezuela	0.522***	30.89	0.441***	15.33	3.70***	53.12	-9.32***	-4.46
Indonesia	0.536***	34.58	0.103***	3.57	2.50***	8.96	-0.61	-0.47
Turkey	0.712***	43.33	0.403***	13.97	2.44***	16.79	-0.78	-0.60
Russia	0.757***	42.85	0.377***	11.96	1.90***	19.16	-0.96	-0.86
Average	0.18	.	0.05	.	2.03	.	-2.03	
Std dev	0.19	.	0.11	.	0.67	.	2.49	
Var(PHB)/var(THB)	.	.	0.34	.	.	.	13.81	
Min	0.003	3.22	-0.059	-3.33	0.320	1.90	-9.32	-4.79
Quartile 1	0.046	10.10	-0.010	-0.43	1.540	14.87	-3.44	-3.03
Quartile 2	0.091	15.54	0.001	0.03	1.990	31.36	-1.74	-1.83
Quartile 3	0.273	22.12	0.072	2.87	2.475	51.79	-0.56	-0.54
Max	0.757	43.33	0.441	15.33	3.700	213.56	5.05	4.47

Key This table shows the total home bias (THB) and the pure home bias (PHB) for the covar-W and logratio measure. Covariances for covar-W are estimated based on monthly returns from January 1994 to December 2012. For covar-W, the THB t-test is the Newey-West significance test in a regression of r_i on $r_{p_i} - r_w$, and for log ratio it is a Newey-West test on the time-series mean of $\log eq_{ij}/m_j$. For both PHB-s, finally, t is the gap between total home bias and the intercept of the foreign-bias regression, scaled by the standard error of that intercept. The logratio measure is the average of annual biases from 2001 to 2012, and its significance test is a Newey-West t test in a regression with just a constant.

Figure 2: **Cross-plots for log-ratio- and covariance-based bias measures: total bias (left) or pure home bias (right)**



tively, the median as a summary measure, the typical gap is 9.1%. Even among the developed markets the typical value is about 4.5% (Q1). Thus, the low U.S. number does not mean that home bias is not a serious problem generally: the U.S., with its well-diversified and large supply of home stocks, is not representative. Statistically, lastly, the evidence is overwhelming across the board, including even for the U.S.[12]

In contrast, the log-ratio figures for total home bias are harder to interpret economically. There is a relation with development, but it is much noisier than for the covar-W-based counterparts.

All this is about total home bias. Table 3 continues the comparison by giving a preview of our results on pure home bias. For the covar-W measure, these are statistically insignificant in the developed markets. In the cross-section of total home bias, two-thirds of the variation is explained by the intercept of the foreign-bias regression, which is our estimate of a normal home bias level in line with the pattern of foreign biases. For the log-ratio measure, the results for pure home bias make less sense. Many countries have a very large negative pure home bias, which implausibly suggests that their citizens strongly dislike their home country compared to a putative zero-distance foreign country. And the correlation with the total home bias is negative: the story is that very home-biased countries actually tend to more strongly *dis*like home, after taking into account distance. The cross-sectional variance of pure home bias is

[12]The t-ratio is a robust t-statistic from the regression of r_i on $r_{p_i} - r_w$.

even 13.81 times that of total home bias, all of which is hard to believe.

In summary, on economic and statistical grounds, covar-W emerges as a much more meaningful measure than log-ratio. We provide a more detailed discussion of the results regarding pure home bias in Section 3.4.

3.3 Foreign-bias regressions: Constructing the foreign bias measures

We now turn to the empirical results for the foreign-bias regressions. For completeness, we present these results for all four measures.

3.3.1 Details about the estimation

The horse-race regressions were run as follows. The data structure corresponds to a three-dimensional panel with 12 years of observations for $N = 41$ home countries and $N - 1 = 40$ destination markets.[13] In practice, home bias is hardly moving over time, an observation that is confirmed by the literature (for example in Bekaert and Wang, 2009; Cooper, Sercu and Vanpée, 2013; and Levy and Levy, 2014). Similarly, almost all of the explanatory variables are highly persistent.[14] As a result, we lose very little information if we simplify the estimation to a two-dimensional problem by just using (41×40) time series averages. We do so for most of our work, as follows.

For the log-ratio measure and the traditional measure, we can immediately calculate the average foreign bias for each home and destination pair over our sample period of 12 years. For the covariance measure, we use a constant matrix, based on 1994–2012 data. [15] The independent variables are all averaged over the twelve year period 2001–2012, except for geographical distance and the common language dummy, which are time-invariant. For some variables, such as the proportion of foreign listings, we have data over a shorter period only; we then use the

[13]One country in our 42-large set, the Philippines, drops out because it was almost completely sealed off from the rest of the world. See Appendix A.

[14]See the online Appendix.

[15]To be in line with the averaging, we should have averaged 12 matrices, one based on 1994–2001, the next on 1995–2002, etc. But the result would have given a lot of weight to the year 2001 (the dotcom crash, accounting scandals, 9/11), and less and less weight, linearly, to data further away from 2001. The final result is not meaningfully affected, but we prefer the equal-weight version.

averages over what we have. The data availability period for each variable is shown in Table 2—

The risk in integrating-out the time dimension is that we may overlook some potentially relevant changes, like the introduction of the euro. In Section 3.5, therefore, among other robustness tests, we do allow distance aversion and pure home bias to change over time by estimating them over two subperiods, and we check whether the estimated changes are related to shifts in the explanatory variables. There are a few such effects, and they are in line with our conclusions.

The equation to be estimated is (6),

$$\begin{aligned} Y_{ij} &= a_i + b_i \overline{X}_{ij} + e_{ij}, \;\; j \neq i, \\ &\quad \text{where } \overline{X}_{ij} := \sum_{k=1}^{K} w_k X_{ijk} \;\; \text{for } j = 1, \cdots, N, \;\text{with } \sum_{k=1}^{K} w_k = 1. \end{aligned} \tag{11}$$

The candidate distance variables are all standardised, which should facilitate the interpretation of the weights within the distance index if, at least, all weights are positive. We correct the standard errors for clustering across the country-pairs in the panel we estimate.

In light of the multicollinearity issues discussed before, we select a best reduced set of variables in a stepwise manner. For each variable, the stepwise selection method calculates the F-statistic that reflects the variable's contribution to the model if it is included. Variables are added one by one to the model, and the statistic for a variable to be added must be significant at a predefined entry level (set at 0.50). After a variable is added, however, the stepwise method reconsiders all the variables already included in the model and deletes any variable that does not produce an F-statistic significant at a predefined stay-in level (set at 0.10). Only after this check is made and the necessary deletions are accomplished can another variable be added to the model. The stepwise process ends when none of the variables outside the model has an F-statistic significant at the entry level and every variable in the model is significant at the stay-in level, or when the variable to be added to the model is the one just deleted from it. The selected distance measures depend on the respective foreign bias measure.

Table 4: **Foreign-bias regression for the competing investment bias measures: Summary figures**

	log-ratio		traditional		covar-W		covar - L	
	Panel A: weights w_k inside the synthetic distance index							
	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat
Tax-capgain$_i j$	.	.	0.121***	5.81	0.562***	5.70	0.004	0.20
Tax-div$_i j$	0.033***	2.65	0.159***	9.37	.	.	0.059***	3.91
Capin$_i j$	-0.012	-0.77	0.102***	4.34	0.677***	6.10	0.097***	5.14
Dist$_i j$	0.287***	12.01	-0.315***	-8.58	0.981***	5.31	.	
D-trade$_i j$	0.298***	9.19	0.308***	8.56	0.399***	2.61	-0.031	-1.61
Language$_i j$	.	.	.	.	.	.	0.092***	6.22
Currency$_i j$	.	.	-0.142***	-6.35	.	.	0.049***	3.07
Flist$_i j$	0.119***	3.86	.	.	.	.	.	.
Size$_i j$	0.107***	5.86	0.746***	21.7	-0.555***	-3.06	0.055***	2.79
Gov$_i j$	0.168***	7.88	0.020	0.69	-1.065***	-3.60	0.674***	19.53
Sum	1.000	.	1.000	.	1.000	.	1.000	.
	Panel B: other regression statistics							
avg b	-0.767	.	0.060	.	-0.003	.	0.085	.
std b	0.420	.	0.014	.	0.002	.	0.011	.
std/avg b	0.548	.	0.237	.	0.780	.	0.128	.
R^2 corrected	0.50	.	0.50	.	0.58	.	0.52	.

Key This table shows the nonlinear least squares estimation results for the equation $Y_{ij} = a_i + b_i[\sum w_k X_{ijk}] + \nu_{ij}$. The left-hand side variables are, respectively, the log-ratio measure, the traditional foreign bias, the covar-W measure, and the covar-L measure. The top panel of the table summarizes the estimation results for the country-specific slope coefficients and shows the estimated weights for the distance variables. 'Average' and 'stdev' refer to the cross-section across 41 country-specific estimates. Significance at the 90%, 95% and 99% confidence levels are indicated with *, ** and ***, respectively. The model is estimated using cross-section clustered standard errors.

3.3.2 Main results

In Table 4, the focus is on the estimated weights w_k for each distance vector $X_{..k}$ within the overall index $\overline{X}_{..}$ (Panel A); regarding the estimated distance aversion coefficients b_i, the table just provides the means and standard deviations (Panel B). We provide that information for each of the four foreign-bias measures, including the two that we discount on validity grounds. For the covar-W and log-ratio measures, which do not suffer from the validity problem uncovered in Section 3.2.1, we then show more details on the slopes, *i.e.* the distance aversions, in Table 5.

To gauge the adequacy of the models on empirical grounds, we first consider purely statistical features, namely the R^2 and the variability of the country-specific slopes. Note, first, that various R^2s are easily interpreted only when there is a common dependent variable and

Table 5: **Foreign-bias regression for the competing investment bias measures: Distance aversions per country**

	covar-W		log-ratio			covar-W		log-ratio	
	estim	t-stat	estim	t-stat		estim	t-stat	estim	t-stat
US	-0.0003	-0.44	-0.184**	-2.12	Egypt	-0.0025***	-2.86	-1.115***	-5.37
UK	-0.0004	-0.51	-0.275***	-3.09	Venezuela	-0.0026***	-3.04	-2.218***	-11.36
Switzerland	-0.0004	-0.57	-0.421***	-4.67	India	-0.0028***	-3.14	-0.454**	-2.14
Netherlands	-0.0005	-0.70	-0.326***	-4.36	Chile	-0.0029***	-3.46	-0.648***	-3.11
Belgium	-0.0008	-1.08	-0.425***	-5.83	Safrica	-0.0029***	-3.16	-0.987***	-3.75
Denmark	-0.0009	-1.11	-0.522***	-5.05	Hong Kong	-0.0029***	-3.56	-0.521***	-4.06
Israel	-0.0009	-1.27	-1.090***	-6.04	Czech Rep	-0.0032***	-3.24	-0.872***	-8.17
Germany	-0.0009	-1.47	-0.422***	-6.63	Mexico	-0.0032***	-3.34	-0.025	-0.18
France	-0.0010	-1.36	-0.503***	-6.45	Singapore	-0.0033***	-4.00	-0.944***	-8.97
Portugal	-0.0011	-1.28	-0.830***	-6.17	Argentina	-0.0035***	-3.65	-1.382***	-6.38
Canada	-0.0011	-1.45	-0.337***	-2.93	Brazil	-0.0035***	-3.66	-1.249***	-6.33
Australia	-0.0011*	-1.82	-0.814***	-6.18	Colombia	-0.0039***	-3.77	-0.685***	-3.40
Italy	-0.0012	-1.58	-0.567***	-5.29	Korea	-0.0042***	-3.93	-0.407***	-2.65
Sweden	-0.0012	-1.49	-0.718***	-7.61	Poland	-0.0043***	-3.70	-0.711***	-6.92
Japan	-0.0013**	-2.06	-0.271**	-2.27	Hungary	-0.0051***	-3.94	-1.011***	-8.6
Spain	-0.0013*	-1.67	-0.889***	-7.46	Malaysia	-0.0053***	-4.39	-1.013***	-8.07
Norway	-0.0013	-1.50	-0.754***	-7.23	Turkey	-0.0055***	-3.99	-0.868***	-5.69
Finland	-0.0014*	-1.68	-1.072***	-8.53	Russia	-0.0072***	-4.03	-0.787***	-5.90
Austria	-0.0014*	-1.89	-0.634***	-7.14	Thailand	-0.0073***	-4.54	-0.705***	-4.87
New Zealand	-0.0014**	-2.12	-1.791***	-9.58	Indonesia	-0.0083***	-4.63	-0.956***	-6.36
Greece	-0.0022**	-2.58	-1.041***	-6.40					

Key This table shows the individual slopes for the equation $Y_{ij} = a_i + b_i[\sum w_k X_{ijk}] + \nu_{ij}$. The left-hand side variables are, respectively, the covar-W and the log-ratio measures. Summary statistics across all 41 countries can be found in the preceding table.

competing sets of regressors, while here we have the opposite configuration.[16] So in our case, the variance explained proves at best an indication. That said, by the R^2 criterion, the covariance measures seem to modestly stand out with an explained variance of 58% for covar-W and 52% for covar-L, followed by the two weight-focused measures (both at 50%).

One could also consider other aspects, like the cross-sectional variation in the slopes b_i: more consensus might seem encouraging. The coefficient of variation is lowest for the covar-L measure and the traditional measure. However, this is simply because the foreign bias regressees for these measures are almost identical for each investing country, so a similar slope

[16] A correction for different dependent variables is available only if the competing Ys are one-to-one functions of each other, which does not apply here. A complicating factor is that the covar-L measure is already a difference between a foreign item and a home one, so that the measure has value zero for home, unlike the other ones. Similarly, converting the other measures into differences would have destroyed the comparability with the prior literature, though, so we abstained. Still, any such transformation might have made the R^2 look quite different.

indicates nothing positive about either the model or the measure.

The two measures which are less dominated by their benchmark, the log-ratio and covar-W, show a high relative variability of slopes between countries. For neither measure is it possible to accept the Null of identical distance aversions b_i for all countries, so in the remainder of the analysis we maintain the assumption of different slopes. The average slopes for these two measures are negative, as we expect, and so are all individual estimates. For the two measures that we rejected for being dominated by the benchmark, the average b is positive, which would be odd if those measures really did measure home bias, cross-sectionally; rather they seem to pick up size ($-m_j$ in the traditional measure) and volatility ($-var(r_{p_j})$, for covar-L) of the host market. Returning to the covar-W and log-ratio measures, we lastly note that even though the signs of the estimated b_is are correct, the correlation between the two sets of distance aversion estimates, 0.22, is low, again suggesting that the estimates measure different phenomena and/or that at least one set is subject to serious estimation noise. One striking difference between the two sets is immediately apparent from Table 5. In that table, the countries are arrayed on the basis of the covar-W distance aversion, which immediately reveals that low estimates of covariance aversion tend to be a characteristic of mature, developed markets: for two-thirds of that group, distance aversion is not even significant.

Returning to Table 4, we next note that the weightings in the distance index are generally well-behaved for the log-ratio and covar-W measures. We see positive weights for the main distance variables (which, combined with a negative b_i, indicate negative effects, i.e. worse foreign bias). We obtain a negative (but insignificant) estimate for the proxy for capital inflow restrictions in the log-ratio regression, and significant negative weights for the governance coefficient the covar-W regression and for relative size in both regressions. As argued in Section 2.3, one cannot have a strong *a priori* view of the sign of the coefficient for these variables: both regressors are examples of $X_{ij} = -X_{ji}$, so that the coefficient picks up the net effect of opposing push and pull considerations.

Both regression procedures select 7 significant weights out of the pool of eleven, and 5 of these are common to both measures. Given the big negative weight for Governance in covar-W and the constraint $\sum_j w_j = 1$, the size of the coefficients is not informative of explanatory power despite the standardisation of the regressors. Two determinants that were often found to be significant in the prior literature, common language and common currency, do not make the final list for either the covar-W or the log-ratio measure.

3.3.3 Non-linearity tests

Since the home bias measure relies on an intercept and that intercept is an extrapolation of the regression line outside the data domain, the linearity of that relation is crucial. We briefly summarize the results for four variants,[17]

$$Y_{ij} = \begin{cases} a_i + b_i\overline{X}_{ij} + c_i[\overline{X}_{ij}]^2 + e_{ij} & \text{(variant 1)} \\ a_i + b_i\overline{X}_{ij} + d_i[\overline{X}_{ij}]^{-1} + e_{ij} & \text{(variant 2)} \\ a_i + b_i\overline{X}_{ij} + c_i[\overline{X}_{ij}]^2 + d_i[\overline{X}_{ij}]^{-1} + e_{ij} & \text{(variant 3)} \\ e^{a_i + b_i\overline{X}_{ij}} + e_{ij} & \text{(variant 4)} \end{cases} \tag{12}$$

For the exponential equation, the estimation did not converge, presumably because of the cross-equation constraint. The results for the other variants provide no good evidence for non-linearities for covar-W: in specifications with country-specific coefficients (c_i for the quadratic term or d_i for the inverse), just two coefficients are significant for the quadratic (Brazil and Turkey), and three for the inverse (Hungary, Turkey and, with just a 10% significance, Brazil). The results are less reassuring for the log-ratio measure, where 8 out of 41 countries, all of them emerging markets, have a significant coefficient for the quadratic term and the same 8 countries also have a significant coefficient for the inverse.

We conclude that for covar-W, non-linearity does not seem to be an issue. For the log-ratio measure there might be a problem for emerging markets. But that measure suffers from other issues, as we have seen, so that we present its results mostly for the sake of completeness.

[17] Tables are available in the online Appendix.

3.4 Pure home bias: Significance and determinants

3.4.1 How large, empirically, is pure home bias?

After our discussion on how to measure distance and how much it matters to whom, we now turn to the resulting pure home bias estimates,

$$PHB_i := Y_{ii} - \hat{a}_i, \tag{13}$$

with $\hat{a}_i$ the home's intercept retrieved from its foreign-bias regression. Table 3, already discussed in Section 3.2.2, shows the estimates and t-ratios for both the total home bias and the pure home bias, once for covar-W and then for the log-ratio. Figure 3 provides a graphical version, showing the two sets of point estimates with their ±2-sigma error zone added, and arranged by increasing pure home bias rather than by total bias, the criterion in Table 3. Unlike in Table 3, the covar-W home biases are no longer multiplied by the relative risk aversion (3), which explains the difference in magnitude between the table and the figure.

From the table, for the covar-W measure, both the mean and the median of the pure bias are much smaller than those for the total home bias. The median is a mere 10 bp, down from 300, and the U.S. figure is even less than 1 bp, down from 10. Out of the 41 estimates, almost half (18) are negative; and only 16, in all, are significant (down from 41/41 for total bias). Some of the right skewness remains, though, and xenophobia remains more widespread than xenophilia: while 23 estimates are positive, a full 14 of these are significant against just two among the 18 negative ones.[18] Also the variability across countries is down, once we correct for distance and distance aversion: the variance of pure home bias is just one-third of the total. We further note that pure and total home bias are positively related: the ranking in Table 3 based on total bias is quite similar to the ranking based on pure bias, in Figure 3, and the Spearman correlation between total and pure hone bias turns out to be 0.85. In line with that last observation, the countries that have low total bias also tend to be the ones with low pure bias, event though, as we have seen, they tend to have a lower distance aversion too. Only Chile (in the low-bias group) and Finland (in the significantly positive group) do not fit the

[18] Singapore and Chile receive a significantly negative pure home-bias score. Neither has an exceptionally low total home bias (Table 3) or a rather high distance aversion (Table 5). Chile, and to a lesser extent Singapore, however get high average scores for distance.

pattern.

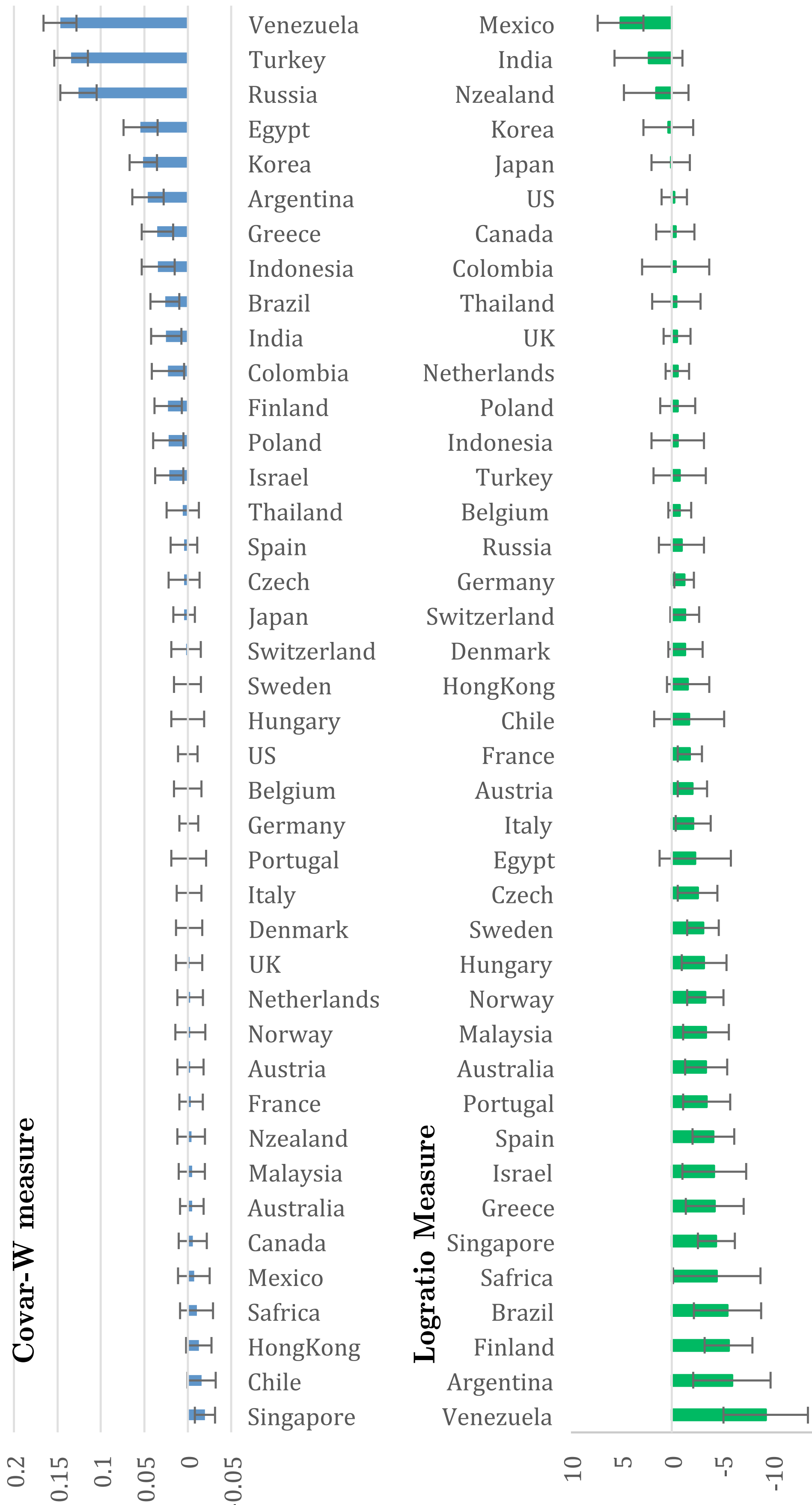
The insignificance of pure home bias for developed economies should not be taken as just a an automatic consequence of the economic insignificance of their total bias. There is no mechanical reason why pure home bias is necessarily a positive fraction of total home bias. Distance aversion, for instance, is country-specific and could have upset the relation. Even for covar-W there are quite a few exceptions where the fraction is negative, for instance. But the absence of a mechanical link is even more vividly illustrated when we turn to the results of the log-ratio measure.

For the log-ratio measure we get a rather confusing picture indeed. Out of the 41 estimated pure home biases, 36 are negative, and the numbers are large in absolute value. The hard-to-believe diagnosis, in short, would be massive xenophilia. Equally puzzling, the correlation between bias before and after the distance correction is negative (–0.31), which, if true, would mean that a higher total home bias would actually hide a strong dislike for the home country, after stripping out distance effects. For the log-ratio, in short, distance does not at all 'explain' all or even part of total home bias: it seems to aggravate it substantially. Nor is there any neat pattern across countries, like the one we observe for covar-W.

The same lack of agreement is illustrated more graphically in Figure 3, where we show the decomposition of home bias into the part explained by foreign bias and the remaining pure home bias. The sum of the bars shows the total home bias. The shaded part is the intercept of the foreign-bias regression, the reflection of distance aversion. The unshaded part is pure home bias. For most cases, in covar-W the positive total bias is associated with a positive intercept in the foreign-bias regression, which then gives us a pure home bias that is much smaller and, in 16 instances, even mildly negative. For the log-ratio measure, in contrast, positive bias tends to come with a very high intercept, leading to a pure home bias that has the opposite sign and unconvincingly exhibits more variability than the original, as we already saw from the variance ratios in Table 3.

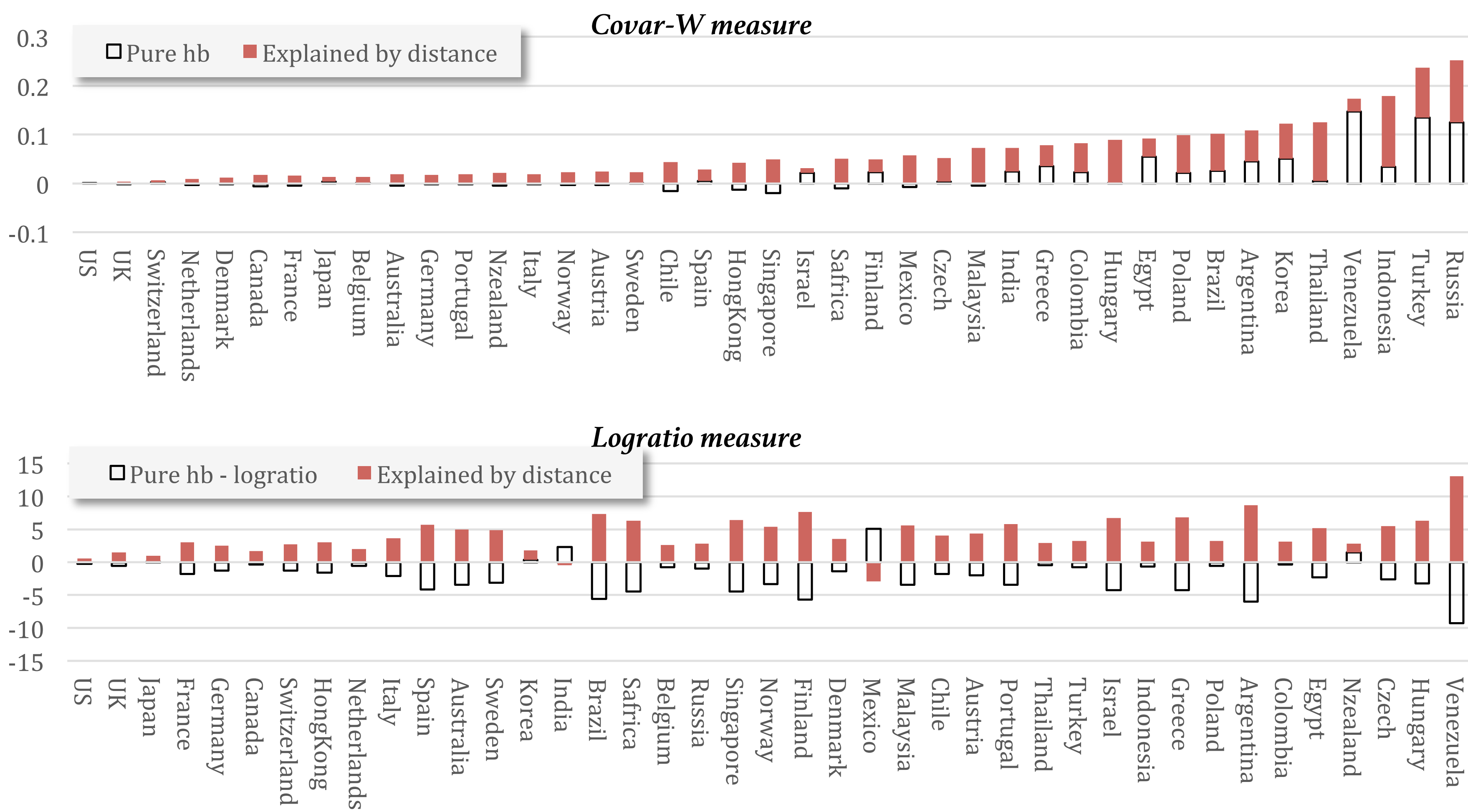
Not surprisingly, then, the log-ratio measure of pure bias shows no relation with the first one, as the rightmost scatter chart in Figure 2 shows. The correlation between both sets of

Figure 3: **Pure home bias resulting from covar-W covariance and log-ratio measures**



Key The figure shows the pure home bias extracted from two possibly valid measures of home bias, the covar-W spread (top) and the log-ratio measure (bottom).

Figure 4: **Decomposing total home bias into its distance-related and residual parts**



Key The figure shows the decomposition of the total home bias into a part explained by distance (shaded) and pure home bias (unshaded) from two possibly valid measures of home bias, the covar-W spread (top) and the log-ratio measure (bottom).

pure home bias estimates, –0.24, is heavily influenced by the lone point right/below, Venezuela: without this item, the correlation is 0.02. Again, either these two pure bias estimates do not measure the same concept, and/or at least one of the two is very noisy, and so the choice of a home bias measure is quite material to the results.

3.4.2 What country characteristics go with strong pure bias and low distance aversion?

The next question is what makes some countries relatively neutral to home and others not. So we take up the pure home bias as calculated from the foreign-bias regressions, and study it via the following second-stage regression:

$$PHB_i = A + DZ_i + u_i, \tag{14}$$

where $PHB_i := Y_{ii} - \hat{a}_i$ (with $\hat{a}_i$ denoting the home's intercept retrieved from its foreign-bias regression), and Z_i a selection of home country specific variables.

Because the variables measuring the level of economic and financial development are all highly correlated, we run 5 alternative regression specifications for which the estimation results are summarized in Table 6. We focus on the covar-W measure; results for the log-ratio measure are provided in Appendix B.

From Panel A in Table 6, the domestic factors associated with pure home bias are capital gains taxes and risk spreads on their government debt: both show up as important in at least three regressions. High levels of either are positively related to pure home bias: that is, low tax rates and low debt spreads at home reduce pure home bias. Low tax rates and low debt spreads are both features that would make a country more attractive to foreigners, not just to residents. Therefore, we view this evidence as being consistent with the pull explanation of pure home bias: foreign investors being keen on country i's assets (which reduces domestic holdings of those assets) rather than home investors being uninterested in the local assets. The regressor 'Forbes' (a relatively high presence of country-i firms in the big-firm league) shows up as significant in one version, and has a similar 'pull' interpretation, with foreigners liking especially the multinationals. In addition, four of the alternative development-related proxies come up as clearly related to pure home bias, with negative signs, meaning that sophisticated

Table 6: **Second-stage regression (covar-W) – Explaining pure home bias and distance aversion**

Panel A: pure home bias

	specification 1		specification 2		specification 3		specification 4		specification 5	
	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat
Intercept	0.146	0.910	0.155	0.990	0.085	0.600	0.034	0.220	0.085	0.70
Tax-capgain	0.648***	5.38	0.638***	5.350	0.548	4.360	0.686***	5.440	0.530***	5.03
Tax-div	0.004	0.11	0.006	0.140	-0.023	-0.640	-0.010	-0.260	0.009	0.28
Size	0.003	0.26	0.003	0.280	0.001	0.100	0.002	0.200	0.006	0.75
Remote	-0.012	-0.46	-0.011	-0.430	0.004	0.140	-0.009	-0.340	-0.023	-1.08
Patriot	-0.009	-0.52	-0.009	-0.510	-0.002	-0.140	-0.001	-0.050	-0.004	-0.27
Forbes	-0.023	-1.57	-0.017	-1.060	-0.009	-0.520	-0.032**	-2.090	-0.005	-0.35
Spread	0.002**	2.51	0.002**	2.360	0.001	0.840	0.002**	2.710	0.000	0.81
HDI	-0.106*	-1.74	.	.	.	.	.	.	.	.
GDP/capita	.	.	-0.026*	-1.98	.	.	.	.	.	.
Soph	.	.	.	.	-0.017**	-2.550	.	.	.	.
Capout	.	.	.	.	.	.	0.002	0.160	.	.
$Rating^2$	.	.	.	.	.	.	.	.	-0.000***	-4.26
Adj R^2	0.498	.	0.511	.	0.543	.	0.451	.	0.650	.

Panel B: distance aversion

	specification 1		specification 2		specification 3		specification 4		specification 5	
	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat
Intercept	-0.025***	-2.83	-0.027***	-3.22	-0.017**	-2.19	-0.019**	-2.38	-0.016**	-2.12
Tax-capgain	-0.011	-1.68	-0.010	-1.60	-0.005	-0.65	-0.011	-1.64	-0.007	-0.99
Tax-div	0.002	0.82	0.002	0.78	0.004**	2.05	0.003	1.24	0.002	1.09
Size	0.001	1.22	0.001	1.24	0.001	1.48	0.001	1.59	0.000	0.90
Remote	0.000	-0.21	0.000	-0.28	-0.002	-1.08	0.000	0.15	0.000	0.16
Patriot	0.002	1.68	0.002*	1.81	0.001	0.97	0.002	1.54	0.001	1.04
Forbes	0.002*	1.96	0.001	1.06	0.001	0.77	0.002**	2.29	0.001	1.17
Spread	-0.000	-0.67	0.000	-0.38	0.000	1.07	0.000	-1.02	0.000	0.78
HDI	0.011***	3.16	.	.	.	.	.	.	.	.
GDP/capita	.	.	0.003***	3.97	.	.	.	.	.	.
Soph	.	.	.	.	0.001***	3.41	.	.	.	.
Capout	.	.	.	.	.	.	-0.003***	-3.26	.	.
$Rating^2$	.	.	.	.	.	.	.	.	0.000***	3.77
Adj R^2	0.448	.	0.514	.	0.469	.	0.456	.	0.498	.

Key Pure home bias (Panel A) and distance aversion (Panel B) are regressed on country characteristics. The last 5 of these are so mutually correlated that they are entered one by one rather than in combination.

countries are less likely to exhibit pure xenophobia.[19] This could be either a 'push' or a 'pull' effect, with development lowering the costs – actual or perceived – of portfolio-investment outflows but also making the home assets attractive to outside investors.

The results so far suggest that for most countries, home bias and foreign bias are not separate phenomena. A single phenomenon, distance aversion, plays a large part in explaining both, as long as the measure of portfolio bias is the covar-W measure. This leaves unanswered the central conundrum of the foreign bias literature: when investors trade off diversification gains against distance, is distance a proxy for a genuine cost, or is it simply picking up a behavioural variable which mirrors no tangible sacrifice? Although we cannot answer this question, our analysis can shed light on one aspect of such behaviour, the investor's sensitivity to differences. Our slope coefficient b_i measures the distance aversion of various countries measured as the sacrifice in diversification relative to a standard distance index. This measure of distance aversion varies across countries, we know, and may very well be zero for many mature economies. So we now examine what variables are related to the different degrees of distance aversion of the individual countries.

Panel B in Table 6 summarizes results from a regression of $\hat{b}_i$ (the slope coefficient of the foreign-bias regression) on country characteristics. It has the same structure as the regressions of pure home bias in Panel A. For the covar-W measure, the only variables which are reliably related to the slope are the alternative measures of development, the last variable in each regression: in all specifications this shows up as significantly positive. So higher development at home makes this slope less negative, that is, higher development comes with lower distance aversion. This is consistent with investors from more developed markets suffering less from a behavioural bias, if the explanation for distance aversion is a behavioural one, or otherwise from lower direct and information costs.

3.4.3 Economic cost of underdiversification

In the above, we motivated the covar-W home bias measure as the deviation between a country's implied expectation and a world average expectation. Another important issue regarding

[19]The square of the rating number is more significant than the spread, and subsumes it.

international portfolio bias is the question, raised by Lévy (2013), of the size of the economic cost associated with the portfolio bias. Lévy measures this cost as the difference between the variance of the actual portfolio and a fully diversified global portfolio. Since our measure of portfolio bias is expressed in terms of the difference between covariances with the actual and the world market portfolio, we are able to relate the estimated parameters directly to Lévy's proposed measure of economic cost.

Lévy's measure of economic cost is the difference in variance between the county's actual portfolio and the world market portfolio:

$$\text{variance gap} := \text{var}(r_{p_i}) - \text{var}(r_w). \tag{15}$$

This can be converted into an equivalent expected-return gap by multiplying by the risk aversion coefficient. One way to decompose this variance spread is

$$\text{var}(r_{p_i}) - \text{var}(r_w) = \sum_j eq_{ij} \underbrace{[\text{cov}(r_j, r_{p_i}) - \text{cov}(r_j, r_w)]}_{\text{covar-W gap}} + \sum_j (eq_{ij} - m_j)\,\text{cov}(r_j, r_w). \tag{16}$$

Thus, Lévy's cost is a weighted average of the covar-W measures, complemented by a portfolio-gap weighted covariance risk.

Lévy focuses on one example, the economic cost of home bias for the U.S., and finds it to be very low, equivalent to 0.12% per annum. We concur: our pure home bias number for that country is about zero. However, we recall from Table 3 that the U.S. already has an unusually low covar-W measure of home bias, the lowest in the set, equal to 0.3%. Other countries, including most developed ones, have much higher measures: France 3.7%, Germany 4.8%, and Italy 5.3%. The comparative lack of potential diversification for the U.S. is a consequence of the large size and diversity of its own stock market.

3.5 Robustness

We checked the robustness of the procedure by examining various alternative specifications for the individual variables and for the functional form. We have already discussed the nonlinear versions of the basic equation, and concluded that the linear equation does well for the covar-W measure, the one we retain in the end on other grounds.

In light of the unexpected sign for the governance coefficient, we also introduced the absolute value of the governance distance, instead of the algebraic value: pure unfamiliarity may be better captured by an un-signed distance. Empirically, it does not: all we obtained was a reduced R-squared. Estimating covariances using various more complex GARCH procedures, lastly, did not qualitatively change the results.

As described in the methodology section, in the above calculations we reduce the three-dimensional panel (42 origins, 41 destinations, 12 years) to a two-dimensional one by averaging over time. To test whether we missed important changes over time, we have re-estimated the model twice: first using averaging over the first four years of data, and then averaging over the last four years. Lastly, we checked whether estimated changes in distance aversion and pure home bias over time are related to changes in the averages for the regressors between the early and late periods. As discussed above, many of the variables are so persistent that the time dimension does not generate nearly as much variation as the cross-sectional differences, so the results can only be suggestive. However, as documented in Appendix B, we find that the reduction in distance aversion is significantly related to an increase in development, consistent with the finding in the main regressions.

As suggested by a referee, we also checked what happens if “distance” is narrowed down to geographical distance. The results, summed up in Table 7, are reassuring in two quite different ways. First, cutting all regressors except geographical distance reduces the $\overline{R}^2$s in the foreign-bias regressions substantially, suggesting we miss important determinants. This is very dramatic for the measures we reject (traditional and covar-L), still quite bad for our favorite, and least damaging for the logratio measure. Yet the pure home bias estimates resulting from both distance metrics are still related, and most remarkably so for covar-W, where the estimates for the full and reduced model, respectively, have similar means, medians and standard deviations, and a correlation of 0.92. Yet this may just have serendipity: using the other measures, the correlations between the full- and reduced-version estimates is much lower. So we do not recommend to pare down the model to just distance, in new studies of portfolio bias.

Table 7: Summary of results when "distance" is reduced to geographical distance

		Log ratio	Traditnl	covar-W	covar-L
foreign-bias regression	Adj R^2, distance only	0.36	0.01	0.14	0.01
	Adj R^2, full model	0.50	0.50	0.58	0.52
pure home bias	Average, full model	-2.026		0.017	
	Average, distance only	0.589	-	0.024	-
	Median, full model	-1.740	-	0.000	
	Median, distance only	0.377		0.003	-
	Stdev, full model	2.489	-	0.038	-
	Stdev, distance only	1.869	-	0.059	-
	Correlation	0.62	-	0.90	-

4 Conclusions

We have developed a measure of pure home bias, defined as that part of total home bias that goes beyond what can be predicted on the basis of the empirical link between the foreign biases and the corresponding distances from the home country.

We find that the most satisfactory way to model that relation between foreign-bias and distance is based on covariances which, unlike portfolio holdings, take into account the covariance structure of returns. Thus, investors do seem to take into account diversification benefits when considering the trade-off with distance. The appropriate benchmark with which to compare their actual portfolios is the global market portfolio. Using this measure, the main variables explaining foreign bias are taxes, restrictions on capital imports in the foreign country, geographical distance, trade intensity, and relative market size. The effects of a shared language and a common currency, two recurrent variables in the prior literature, seem to be subsumed by the covariances. Countries differ in their degree of distance aversion, with the least distance averse being the most developed. For developed markets, which form a majority of the 41 countries we study, pure home bias is insignificantly different from zero, and in general 70% of the variation in total home bias is explained by the foreign bias pattern.

For almost all developed markets, then, the home country is very much like a foreign country with zero distance. Investors do not appear to exhibit a pure fear of foreign investment separate from their general dislike of distance. This does not mean, of course, that home bias does not exist. Rather, it means that home bias and foreign bias collapse into a single puzzle: why do investors allocate lower portfolio weights to more distant countries? Specifically, they give up diversification by avoiding distant markets but there is no systematic benefit to this behaviour that has yet been identified (as far as we know). We actually find distance aversion to be correlated with lack of development, and this is not what one would expect if all relevant costs were objective out-of-pocket expenses and all investors well-informed and rational.

For a smaller part of the sample, all of them riskier and newer markets, pure home bias measured as the expectations gap implied by the observed covariance gap ranges up to 15% per annum, even 40% for the three countries that are virtually walled off from the rest of the world.[20] We identify the country characteristics that are associated with the degree of that pure home bias. These are mainly a high tax rate and a poor credit standing. A strong pure home bias in high-tax/low-reputation countries presumably reflects the unwillingness of foreigners to invest in such markets, not the locals' eagerness to do so.

References

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

Bekaert, G., and C.R. Harvey (1997). Emerging equity market volatility, Journal of Financial Economics 43(1), 29–77.

Bekaert, G., C.B. Erb, C.R. Harvey, and T.E. Viskanta (1997). What matters for emerging equity market investments, Emerging Markets Quarterly 1, 17–46.

Bekaert, G., C.B. Erb, C.R. Harvey, and T.E. Viskanta (1998). Distributional characteristics of emerging market returns and asset allocation. Journal of Portfolio Management 24(2), 102–116.

Bekaert, G., and Wang, X. (2009). Home bias revisited. Available at SSRN: http://ssrn.com/abstract=1344880.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, Review of Economics and Statistics 73(2), 498–505.

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

Chin, M.D., and H. Ito (2006). What Matters for Financial Development? Capital Controls, Institutions, and Interactions, Journal of Development Economics 81(1), 163–192.

Coelho, B., and K.P. Gallagher (2010). Capital Controls and 21st Century Financial Crises: Evidence from Colombia and Thailand, PERI Working Paper Series, WP 213, University of Massachusetts at Amherst.

[20]This includes shadow costs: if Russia's restrictions were abolished, Russians would require 40% extra to hold on to their almost fully home-biased portfolio despite its high variance.

Coeurdacier, N., and S. Gibaud (2011). International portfolio diversification is better than you think, Journal of International Money and Finance 30(2), 289–308.

Cooper, I., and E. Kaplanis (1986). Costs to crossborder investment and international equity market equilibrium, in Recent Developments in Corporate Finance, edited by J. Edwards, J. Franks, C. Mayer, and S. Schaefer, Cambridge University Press, pp. 209–240.

Cooper, I., and E. Kaplanis (1994). Home bias in equity portfolios, inflation hedging, and international capital market equilibrium. Review of Financial Studies, 7(1), 4560.

Cooper, I., P. Sercu, and R. Vanpée (2013). The equity home bias: Survey and synthesis. Foundations and Trends in Finance, 7(4), 289–416.

Fernández, A., A. Rebucci, and M. Uribe (2014). Are capital controls prudential? An empirical investigation, NBER Working Paper No. 19671.

French, K.R., and J.M. Poterba (1991). Investor Diversification and International Equity Markets. American Economic Review, 81(2), 222226.

Friend, I., and M.E. Blume (1975). The Demand for Risky Assets. The American Economic Review 65(5), 900–922

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

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

Karlsson, A., and L. Nordén (2007). Home sweet home: Home bias and international diversification among individual investors. Journal of Banking & Finance 31(2), 317–333.

Ke, D., L. Ng, and Q. Wang (2010). Home bias in foreign investment decisions. Journal of International Business Studies 41(6), 960–979.

Kho B.-C., R.M. Stulz, and F.E. Warnock (2009). Financial Globalization, Governance and the Evolution of the Home Bias, Journal of Accounting Research 47(2), 597–635.

Kilka, M., and M. Weber (2000). Home Bias in International Stock Return Expectations. Journal of Psychology and Financial Markets 1(3–4), 176–192.

Kohers, G., N. Kohers, and T. Kohers (2006). The risk and return characteristics of developed and emerging stock markets: The recent evidence. Applied Economic Letters 13, 737–743.

Lau, S.T., L. Ng, and B. Zhang (2010). The world price of home bias, Journal of Financial Economics 97(2), 191–217.

Levy H., and M. Levy (2014). The home bias is here to stay, Journal of Banking and Finance 47, 29–40.

Levy H. (2013). What is the economic cost of the investment home bias? Available at SSRN: http://ssrn.com/abstract=2345193

Longin, F., and B. Solnik (1995). Is the correlation in international equity returns constant: 1960–1990?, Journal of International Money and Finance 14(1), 3–26.

Mishra, A.V., and R.A. Ratti (2013). Home bias and cross border taxation, Journal of International Money and Finance 32, 169–193.

Morse, A., and S. Shive (2011). Patriotism in your portfolio. Journal of Financial Markets 14(2), 411–440.

Portes, R., and H. Rey (2005). The determinants of cross-border equity flows. Journal of International Economics, 65(2), 269296.

Saborowski C., S. Sanya, H. Weisfeld, and J. Yepez (2014). Effectiveness of Capital Outflow Restrictions, IMF Working Paper WP/14/8.

Sarkissian, S., and M.J. Schill (2004). The overseas listing decision: New evidence of proximity preference, Review of Financial Studies, 17(3), 769–810.

Sarkissian, S., and M.J. Schill (2009). Are there permanent valuation gains to overseas listing?, Review of Financial Studies, 22(1), 371–412.

Sarkissian, S., and M.J. Schill (2014). Cross-Listing Waves, Journal of Financial and Quantitative Analysis, forthcoming.

Schindler, M. (2009). Measuring financial integration: A new data set, IMF Staff Papers 56(1), 222–238.

Seasholes, M.S., and N. Zhu (2010). Individual investors and local bias, The Journal of Finance 65(5), 1987–2010.

Sercu, P., and R. Vanpée (2008). Estimating the Costs of International Equity Investments, Review of Finance, 12(4), 587–634.

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

Appendices

A Details on the regressors

A.1 Variables related to foreign bias

For foreign-bias regressions the variables are measured in such a way that the home country has zero distance for each variable and an increasing value of the variable is expected to decrease investment in that equity market.

A.1.1 Explicit costs and barriers to capital inflows

Taxes are an explicit cost that has been found to be significant in many studies of foreign bias. The tax variable enters into our regression as a differential rate, notably the difference between the taxes on returns from asset j paid by investor i and those paid by a local (j), i.e. $Tax_t^{ij} - Tax_t^{jj}$. We have separate regressors for taxes on dividends ($Tax - div_{ij}$) and on capital gains ($Tax - capgain_{ij}$).[21] Tax rates on capital gains were kindly provided to us by Anil Mishra, and dividend withholding tax rates were obtained from PriceWaterhouseCoopers' Worldwide.[22]

An underinvestment bias towards a particular foreign country may be caused by specific restrictions on capital inflows to that country. Based on the IMF's *Annual Report on Exchange Arrangements and Exchange Restrictions* (*AREAER*), Schindler (2009) developed detailed indices for capital controls that allow distinguishing between inflows *vs* outflows. Data on 91 countries from 1995 to 2005 from his 2009 flagship paper are publicly available, and were recently updated to 2012 by Fernandéz, Rebucci and Uribe (2014). We use Schindler's subindex for capital inflow restrictions (*Capin*) to measure the impact of controls on incoming capital on the bias towards country j.

Based on the index of capital controls, it can be argued that some countries have to be excluded from our sample as either host or home markets for specific years. We want to model the relation between a foreign investment bias and the distance between two countries. However, if a country forbids its citizens to invest abroad with a binding restriction, the logical consequence is that this country has a maximal home bias which is completely unrelated to distance. Hence, we exclude from being a home country those countries that had a maximum value for the index of capital outflow controls. Only one country, the Philippines, had extreme capital outflow controls over our full sample period, thus reducing our initial sample of 42 countries to 41. Other emerging markets like Russia and India are excluded for specific years only. Following a similar reasoning, if a country has extreme controls on capital inflows, foreigners will not buy this country's shares, even if the country is close to the home market.

---

[21] So, unlike Mishra and Ratti (2013), for instance, we do not pre-multiply the tax rate by the taxable return. One reason is that then it is no longer clear whether the variable stands for returns or for taxes. An additional source of worry with including returns is that the taxable basis for capital gains is unclear. Some countries have no capital gains taxes at all, others have a reduced tax rate, and still others use the full rate; in addition, losses can be deducted in some countries but not in all, and loss carry-forward rules are equally different. Finally, if the taxable basis is included, it should be as an expected capital gain, not a realized one. So we let the regression coefficient pick its own implied average taxable basis.

[22] For most countries, the tax rate withheld on dividends, $\tau^{ij}_{d,t}$, differs, depending on the corresponding foreign country. For example, in Denmark, the general withholding tax rate on dividends paid to non-residents is 28% in 2010. However, Denmark has a tax treaty with all member countries of the EU, for whom the tax rate is reduced to 15%, with exception for Greek residents, for which Denmark withholds 18% of the gross dividends.

Table A1: **Countries excluded from the sample due to strict restrictions on capital flows**

		Argentina	Brazil	India	Malaysia	Philippines	Poland	Russia	Thailand	Venezuela
2001	Inflow									
	Outflow				X	X		X		
2002	Inflow							X		
	Outflow				X	X		X		
2003	Inflow							X		
	Outflow	X				X		X		
2004	Inflow									
	Outflow	X		X	X	X		X	X	
2005	Inflow							X		
	Outflow	X		X	X	X	X	X	X	
2006	Inflow							X		
	Outflow			X	X	X	X	X	X	
2007	Inflow							X		
	Outflow			X	X	X	X	X	X	
2008	Inflow									X
	Outflow	X		X	X	X	X		X	
2009	Inflow		X							X
	Outflow			X	X	X	X		X	
2010	Inflow									X
	Outflow		X	X	X	X	X			
2011	Inflow			X						X
	Outflow		X	X		X	X			
2012	Inflow									X
	Outflow		X	X		X	X			

Therefore, we exclude countries as destinations if they have a maximum value for the index of capital inflows. Very few countries forbid capital inflows, though. Only Russia and Venezuela do, and even only for a few years. Table A1 shows for each year all the home and destination countries that are excluded from our sample.

A.1.2 Familiarity indicators

We employ 6 indicators that proxy for familiarity between two countries: the geographical distance, a trade-based distance, a common language indicator, a common currency indicator, a cross-listing ratio, and the relative size of the country. Each of these variables are discussed in turn.

The first familiarity measure is the geographical distance ($Dist_ij$) between the home and destination country, calculated following the great circle formula using latitudes and longitudes of the most important city (in terms of population) or of a country's capital. Bilateral distances are obtained from http://www.cepii.fr/.

The second variable captures the separateness between countries through trade ($D\text{-}trade_i j$). The trade-separateness variable is constructed as follows. Denote the value of exports from country i to country j by X_{ij}. We want to scale actual bilateral trade, $X_{ij}+X_{ji}$, by a maximum level that is achievable given the countries' sizes. Denote i's share in combined absorption by ζ_i, and denote the national outputs by P_i and P_j. In a perfectly frictionless and integrated world, i would export $(1-\zeta_i)\, P_i$ to j and import $\zeta_i\, P_j$ from j, so total trade would equal

$$\hat{X}_{ij} + \hat{X}_{ji} = (1-\zeta_i)\, P_i + \zeta_i\, P_j \tag{17}$$

We use unity minus the square root of the ratio actual/max so that this distance measure becomes zero in the case of trade between i and itself:

$$D\text{-}trade_i j = 1 - \sqrt{\frac{X_{ij} + X_{ji}}{(1-\zeta_i)\, P_i + \zeta_i\, P_j}}\,. \tag{18}$$

The data on exports and imports are from the IMF Direction of Trade Statistics (retrieved via Datastream). The total production is measured by a country's GDP, and aggregate expenditures are calculated by subtracting the current account balance from the GDP. The IMF Financial Statistics report data on the current account balance up to 2008. Current account balances as of 2009 and GDP data were obtained from the World Bank.

Apart from the physical or economic distance, familiarity can also arise between countries that share the same language. Therefore, we introduce a dummy variable that takes the value of unity if two countries have a different language and is equal to zero if the home and host country have a common language ($Language_i j$).

Some research has shown that the equity home bias within the Eurozone has fallen since the introduction of the euro (among others, Coeurdacier and Guibaud, 2011, Baele, Pungulescu and ter Horst, 2007), mainly because Eurozone members benefit from a decrease in implied transaction costs. In general, we expect that countries sharing a currency have lower implicit costs for cross-border transactions. We capture this effect by introducing a dummy variable ($Currency_i j$) that takes the value of zero when the home and host country have a common currency and unity otherwise. Firms that cross-list their shares on a foreign stock market are typically better known in this host market. Therefore, we construct a foreign-listing measure

($Flist_ij$) based on the ratio of the number of cross-listed shares from host country j on the stock market of country i ($Nshare_{ji}$) and the number of shares listed on the domestic stock market ($Nshare_{ii}$):

$$Flist_{ij} := 1 - \sqrt{\frac{Nshare_{ji}}{Nshare_{ii}}}. \tag{19}$$

The data on the number of cross-listed shares for each home and host country were obtained from Sarkissian and Schill (2014), who conducted a survey of the foreign listings on stock exchanges from 73 home countries at the end of 1998, 2003 and 2006. These survey results update earlier work (Sarkissian and Schill 2004, 2009). The number of domestic stock listings was obtained from the World Federation of Exchanges and the websites of the stock markets of each country.

Size breeds familiarity too. For instance, within a given foreign market, Kang and Stulz (1997) show that foreign investors have a strong preference for large firms. This can be due to the fact that large firms are more visible, which implies that information gathering costs and information asymmetries are lower for bigger firms.[23] We hypothesise that, across countries, the same holds: investors hear more about large countries and are more inclined to invest there. For the pure weight-based measures, we could also add a diversification argument: large markets are better diversified, so its residents will be less inclined to invest abroad. For the covariance measures, of course, one cannot invoke diversification anymore.

We measure the relative size ($Size_ij$) of the foreign country as the log of its GDP:

$$Size_{ij} := \log \frac{GDP_i}{GDP_j}. \tag{20}$$

The GDP is measured in USD and the data were obtained from the World Bank. So in line with the other measures for distance, where less familiar markets get higher values, our size measure is equal to zero when the home country is equal to the host country, positive when the host country j is smaller than the home country i, and negative when the host country's GDP is larger than the home market.

[23] Also, large firms are in general more liquid than small firms, which reduces transaction costs.

A.1.3 Financial market sophistication and governance

Well developed and transparent financial markets should attract more foreign capital than less sophisticated capital markets. Portes and Rey (2005) show that financial market sophistication is positively correlated with cross-border equity holdings. Portes and Rey distinguish between financial market development (measured by private credit over GDP) and financial market sophistication as measured by the World Economic Forum, and show that the latter is better able to explain foreign equity holdings than financial market development. We accordingly use the financial market sophistication index ($Soph_ij$) retrieved from the Financial Development Reports published by the World Economic Forum.

There is compelling evidence that proper governance, transparent policy making, and a low level of political risk, are important factors in attracting international capital flows. Gelos and Wei (2005) show that good governance at the country level and at the company level both positively influence international portfolio holdings, but the effect of government transparency is more pronounced. To capture the effect of governance on international equity holdings (Gov), we use the Government Effectiveness indicator from the World Bank Governance Indicators. To reflect the intuition that poorer governance and a lower level of financial market sophistication create a bigger distance for attracting foreign investments, the two variables are defined as

$$Soph_{ij} := -(Soph_j - Soph_i), \quad (21)$$

$$Gov_{ij} := -(Gov_j - Gov_i). \quad (22)$$

A.2 Variables related to pure home bias

To model pure home bias, we use the explicit costs and border controls in the home country, economic development indicators, the average remoteness of the home country, the patriotic nature of its citizens, and variables related to reputation and creditworthiness. These variables are not in difference form since they describe characteristics of the home country.

A.2.1 Explicit costs and capital outflow controls

Corresponding to the tax rate in the foreign country, we use two domestic tax variables: the domestic withholding tax on dividends ($Tax\text{-}div_i$) and the tax rate on capital gains (Tax-

$capgain_i$). We expect that investors residing in low-tax countries hold more domestic stocks in their portfolios.

Another potential reason for a high level of home bias, which is purely driven by the home market, is that investors are less able to invest abroad due to restrictions on capital outflows. We include the Schindler index measuring the intensity of capital outflow controls (*Capout*) (Schindler, 2009, Ferandez, Rebucci and Uribe, 2014).

A.2.2 Size of the home country

We take the log of the home country's GDP to control for its size. Because larger countries have better diversification opportunities at home, we expect the pure home bias to be positively related to size.

A.2.3 Economic and institutional development

Several studies have found a link with development indicators like governance and political risk, financial market development, and equity home bias. Kho, Stulz and Warnock (2009) show that countries with a poor quality of institutions and poor investor protection exhibit a high equity home bias. The reason, they argue, is that a high level of insider ownership is an optimal response to poor corporate and country level governance.

We include three measures to capture the degree of economic and institutional development: the score of a country on the Human Development Index (HDI_i), the home country's GDP per capita (GDP/cap_i), and the Government Effectiveness indicator (Gov_i), all defined before. We expect each of these development indicators to correlate negatively with the pure home bias.

A.2.4 Remoteness of the home country

If a country is more remote, on average, from the rest of the world, it will in general be more difficult for its citizens to obtain information about foreign countries. We construct a variable that measures the average distance between the home country and all hosts included in the sample ($Remote_i$). We expect this variable to correlate positively with pure home bias.

A.2.5 Patriotism

Besides rational reasons, overweighting the domestic market may also be driven by a behavioural bias. Behavioural drivers for the equity home bias may include over-optimism (Kilka and Weber, 2000), the feeling of familiarity with domestic firms (Ke *et al.*, 2010), overconfidence (Karlsson and Nordén, 2007), or patriotism (Morse and Shive, 2011).

A general problem with behavioural-based constructs is that they are difficult to measure. Especially at an aggregate level, data are scarce. One exception is patriotism. The World Values Survey (WVS) contains the answers to a global questionnaire, conducted in almost 100 countries with almost 400,000 respondents. For each country, the WVS aimed to obtain responses of at least 1,000 people. We use the survey waves for the periods 1999–2004, 2005–2009 and 2010–2014. In line with Morse and Shive (2011), our measure for patriotism ($Patriot_i$) is based on the responses to the question 'How proud are you to be [Nationality]?'. The survey responses are coded from 1 to 4, and we use the country's score average over respondents.

A.2.6 Reputation and other 'pull' factors.

One usually looks for home-country variables that make the local investors more or less reluctant to invest abroad ('push' factors), but there could also 'pull' factors, where foreigners are keen on the country's assets, thus leaving fewer of the locals that one would otherwise have predicted. One such feature is a high value weight of large, essentially international companies in the country's stock market. Examples that come to mind are the Netherlands (with Royal Dutch Shell, ING, Philips, Akzo, Unilever, etc.) or Switzerland (Nestlé, ABB, banks and pharmaceutical companies). Our proxy is the weight of a country's members in the Forbes Global 2000 List, scaled by the country's weight in the global stock market capitalization ($Forbes_i$).[24] This way, a score smaller than unity means that in general, companies are relatively small and local, while a value higher than unity indicates the presence of more large firms in the country's total stock market capitalization.

[24] The Forbes Global 2000 List is an annual ranking of the 2000 largest public companies in the world. The list is published by *Forbes* magazine and is available from http://www.forbes.com/global2000/.

We add two more reputation variables: (i) the sovereign interest rate spread in the home country over the 10-year U.S. treasury rate ($Spread_i$), and (ii) country rating: these are Moody's country ratings, converted to a numerical scale (Aaa=24 to C=4) and then squared to reflect the non-linear impact of a rating ($Rating_i^2$). Sovereign spread for the emerging countries were downloaded from World Bank staff calculations available on Datastream, where available. For other countries, we copied CDS spreads. Moody's ratings are from www.moodys.com.

Finally, we note that tax could also be a pull variable, at least for dividends, which are generally taxed in the source country. A low tax rate could serve to attract foreign investors or a high tax rate to deter them. So although we have listed tax as a push variable above, it could also represent pull. The interpretation will depend on the coefficient and the type of taxation. If a high tax rate on dividends results in a high pure home bias, then it will represent pull in the sense that it deters foreigners. If a high dividend tax rate results in a low pure home bias, then it will represent push in the sense that it causes home investors to invest abroad.

B Appendix: Additional results

We report results for, first, the log-ratio measures (what country characteristics go with high distance aversion and pure home bias) and, second for changes over time (2001–2004 versus 2009–2012), in covariance-based pure home bias and distance aversion. In these last regressions we first run the model using, everywhere, averages over the first four years, then once more using averages over the last four years, and finally we regress changes in the estimated distance aversions and pure home biases to changes in the regressors between 2001–2004 and 2009–2012. The tables shown refer to that final stage.

Table A2: **Second-stage regression (log-ratio) – Explaining pure home bias and distance aversion**

Panel A: Log-ratio-based pure home bias

	specification 1		specification 2		specification 3		specification 4		specification 5	
	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat
Intercept	-16.056	-1.160	-14.564	-1.080	-19.625	-1.560	-20.135	-1.550	-21.945*	-1.72
Tax-capgain	-20.604*	-1.98	-21.527**	-2.090	-24.807**	-2.200	-19.901*	-1.890	-19.009*	-1.72
Tax-div	0.873	0.26	1.072	0.330	-0.457	-0.140	0.320	0.100	0.163	0.05
Size	2.049**	2.36	2.067**	2.410	1.974**	2.290	1.978**	2.240	2.049**	2.31
Remote	-1.109	-0.51	-1.082	-0.500	-0.411	-0.190	-1.210	-0.540	-1.006	-0.45
Patriot	-0.077	-0.05	-0.128	-0.080	0.305	0.210	0.158	0.100	0.387	0.26
Forbes	-0.492	-0.38	-0.020	-0.010	0.114	0.080	-0.734	-0.580	-0.858	-0.61
Spread	-0.089	-1.60	-0.097*	-1.730	-0.132*	-1.930	-0.082	-1.470	-0.084	-1.30
HDI	-5.462	-1.04	.	.	.	.	.	.	.	.
GDP/cap	.	.	-1.538	-1.37	.	.	.	.	.	.
Soph	.	.	.	.	-0.753	-1.280	.	.	.	.
Capout	.	.	.	.	.	.	0.866	0.660	.	.
$Rating^2$	.	.	.	.	.	.	.	.	-0.001	-0.15
Adj R^2	0.120	.	0.141	.	0.134	.	0.102	.	0.091	.

Panel B: Log-ratio-based distance aversion

	specification 1		specification 2		specification 3		specification 4		specification 5	
	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat
Intercept	-3.481**	-2.04	-3.489**	-2.06	-3.934**	-2.51	-3.548**	-2.29	-4.085**	-2.60
Tax-capgain	-4.015***	-3.13	-4.044***	-3.14	-4.261***	-3.02	-4.209***	-3.35	-3.892***	-2.86
Tax-div	-0.545	-1.32	-0.546	-1.33	-0.668	-1.65	-0.562	-1.43	-0.615	-1.50
Size	0.426***	3.96	0.427***	3.98	0.420***	3.89	0.408***	3.88	0.427***	3.90
Remote	-0.432	-1.60	-0.427	-1.58	-0.375	-1.36	-0.489*	-1.83	-0.426	-1.54
Patriot	-0.058	-0.30	-0.052	-0.27	-0.015	-0.08	-0.077	-0.42	-0.011	-0.06
Forbes	0.278*	1.74	0.307*	1.78	0.309	1.68	0.294*	1.94	0.247	1.43
Spread	-0.014**	-2.02	-0.014**	-2.06	-0.017*	-1.97	-0.014**	-2.04	-0.014*	-1.73
HDI	-0.574	-0.88	.	.	.	.	.	.	.	.
GDP/cap	.	.	-0.128	-0.92	.	.	.	.	.	.
Soph	.	.	.	.	-0.056	-0.76	.	.	.	.
Capout	.	.	.	.	.	.	0.246	1.57	.	.
$Rating^2$	.	.	.	.	.	.	.	.	0.000	-0.22
Adj R^2	0.527	.0	528.	0.524	.0	550.	0.516	.	.	.

Key The log-ratio measures of pure home bias (Panel A) and distance aversion (panel B) are regressed on country characteristics. The last 5 of these are so mutually correlated that they are entered one by one rather than in combination.

Table A3: **Determinants of evolution over time of pure home bias and distance aversion**

Panel A: evolution over time of pure home bias

	specification 1		specification 2		specification 3		specification 4		specification 5	
	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat
Intercept	-0.001	-0.02	-0.012	-0.50	-0.012	-0.44	-0.005	-0.21	-0.004	-0.14
Tax-capgain	-0.044	-0.35	-0.005	-0.05	-0.058	-0.46	-0.058	-0.49	-0.059	-0.42
Tax-div	-0.054	-0.47	-0.066	-0.65	-0.048	-0.43	-0.056	-0.53	-0.048	-0.41
Size	-0.075	-0.80	1.031**	2.59	-0.046	-0.56	-0.057	-0.81	-0.070	-0.83
Patriot	-0.249**	-2.36	-0.246***	-2.77	-0.270	-2.73	-0.269***	-2.90	-0.252**	-2.54
Forbes	-0.134	-0.62	-0.087	-0.45	-0.198	-0.89	-0.047	-0.23	-0.131	-0.61
Spread	-0.000	-0.17	0.000	0.26	-0.000	-0.18	0.000	-0.19	-0.000	-0.30
HDI	-0.063	-0.09	.	.	.	.	.			
GDP/cap	.	.	-1.190***	-2.83	.	.	.			
Soph	.	.	.	.	-0.033	-1.03	.	.		
Capout	.	.	.	.	.	.	0.112**	2.26	.	
$Rating^2$	.	.	.	.	.	.	.	.	-0.000	-0.26
Adj R^2	-0.003	.	0.193	.	0.028	.	0.131	.	-0.001	.

Panel B: evolution over time of distance aversion

	specification 1		specification 2		specification 3		specification 4		specification 5	
	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat	coeff	Tstat
Intercept	-0.004***	-2.78	-0.004**	-2.60	-0.004**	-2.53	-0.004**	-2.62	-0.002*	-1.72
Tax-capgain	-0.006	-0.78	-0.006	-0.72	-0.006	-0.78	-0.005	-0.69	0.004	0.52
Tax-div	-0.005	-0.73	-0.005	-0.79	-0.005	-0.76	-0.005	-0.81	-0.009	-1.48
Size	0.012**	2.24	0.022	0.83	0.016***	3.31	0.014***	3.36	0.009**	1.98
Patriot	-0.006	-0.99	-0.004	-0.69	-0.004	-0.75	-0.003	-0.57	-0.005	-0.92
Forbes	-0.015	-1.16	-0.016	-1.23	-0.017	-1.29	-0.021*	-1.73	-0.017	-1.51
Spread	0.000	-1.67	0.000	-1.56	0.000	-1.63	-0.000*	-1.74	0.000	0.67
HDI	0.038	0.95	.	.	.	.				
GDP/cap	.	.	-0.007	-0.24	.	.	.			
Soph	.	.	.	.	-0.001	-0.33	.	.		
Capout	.	.	.	.	.	.	-0.007**	-2.24	.	
$Rating^2$	.	.	.	.	.	.	.	.	0.000***	3.31
Adj R^2	0.302	.	0.284	.	0.285	.	0.377	.	0.461	.

Key The changes, 2001–2004 versus 2009–2012, in pure home bias (Panel A) and distance aversion (Panel B) are regressed on country characteristics. The last 5 of these are so mutually correlated that they are entered one by one rather than in combination.