Information Immobility and the Home Bias Puzzle

Stijn Van Nieuwerburgh and Laura Veldkamp *
New York University
Stern School of Business

November 21, 2007

Abstract

Many papers have argued that home bias arises because home investors can predict payoffs of their home assets more accurately than foreigners can. But why does this information advantage exist in a world where investors can learn foreign information? We model investors who are endowed with a small home information advantage. They can choose what information to learn before they invest in many risky assets. Surprisingly, even when home investors can learn what foreigners know, they choose not to. The reason is that investors profit more from knowing information that others do not know. Allowing investors to learn amplifies their initial information asymmetry. The model explains local and industry bias as well as observed patterns of foreign investments, portfolio out-performance and asset prices. Finally, we outline new avenues for empirical research.

JEL classification: F30, F40, D82.
Keywords: Home bias, asymmetric information, learning.

*Stijn Van Nieuwerburgh: svnieuwe@stern.nyu.edu, NYU Stern, Finance Department, 44 West 4th St., 9th floor, New York, NY 10012. Laura Veldkamp: lveldkam@stern.nyu.edu, NYU Stern, Economics Department, 44 West 4th St., 7th floor, New York, NY 10012. Thanks to Campbell Harvey and an anonymous associate editor and referee for their comments, which substantially improved the paper. Thanks also to Dave Backus, Pierre-Olivier Gourinchas, Urban Jermann, David Lesmond, Karen Lewis, Anthony Lynch, Arzu Ozoguz, Hyun Shin, Chris Sims, Eric Van Wincoop, Mark Wright, participants at the following conferences: AFA, Banque de France conference on Economic Fluctuations Risk and Policy, Budapest SED, CEPR Asset Pricing meetings in Gerzensee, CEPR-National Bank of Belgium conference on international adjustment, Cleveland Fed International Macroeconomics and Finance conference, Econometric Society, EEA, Financial Economics and Accounting, Financial Management Association, Prague workshop in macro theory, NBER EF&G meetings and NBER summer institute in International Finance and Macro and seminar participants at Columbia GSB, Emory, Illinois, Iowa, GWU, LBS, LSE, Minneapolis Fed, MIT, New York Fed, NYU, Ohio State, Princeton, Rutgers, UCLA, UCSD, and Virginia for helpful comments and discussions. Laura Veldkamp thanks Princeton University for their hospitality and financial support through the Peter B. Kenen fellowship.

Observed returns on national equity portfolios suggest substantial benefits from international diversification, yet individuals and institutions in most countries hold modest amounts of foreign equity. Many studies document such home bias (see French and Poterba (1991), Tesar and Werner (1998) and Ahearne, Griever and Warnock (2004)). One hypothesis is that capital is internationally immobile across countries, yet this is belied by the speed and volume of international capital flows among both developed and developing countries. Another hypothesis is that investors have superior access to information about local firms or economic conditions. But this seems to replace the assumption of capital immobility with information immobility. If an American wished, she could obtain information about foreign firms. Such cross-border information flows could potentially undermine the home bias. This criticism of home bias theories undermines many asymmetric-information-based theories in finance and raises the question: When investors can choose what information to learn or what data to collect, can information asymmetry survive?

Most existing models of asymmetric information in financial markets are silent on information choice.[1] A small but growing literature studies how much information investors acquire about one risky asset or models a representative agent who, by definition, cannot have asymmetric information.[2] Instead of asking how much investors learn, we ask which assets they learn about. To answer this question requires a model with three features: information choice, multiple risky assets to learn about, and heterogeneous agents so that information asymmetry is possible.

We develop a two-country general equilibrium, rational expectations model where investors first choose what home or foreign information to acquire, and then choose what assets to hold. The prior information each home investor has about each home asset's payoff is slightly more precise than the prior information foreigners have. The reverse is true for foreign assets. This prior information advantage may reflect what is incidentally observed from the local environment. Investors choose whether to acquire additional information about home assets or to acquire information about foreign assets. The interaction of the information decision and the portfolio decision causes the

[1]Recent work on asymmetric information in financial markets includes Banerjee (2007), Ozdenoren and Yuan (2007) and Yuan (2005). The canonical reference on asymmetric information with multiple assets is Admati (1985). Work on asymmetric information and the home bias, in particular, includes Pastor (2000), Brennan and Cao (1997) and Portes, Rey and Oh (2001).

[2]Recent work on information choice in finance includes Peress (2006) and Dow, Goldstein and Guembel (2007). The canonical references in this literature, Grossman and Stiglitz (1980) and Admati and Pfleiderer (1990), are also about one risky asset. Our paper also differs from Calvo and Mendoza (2000) who argue that more scope for international diversification decreases the value of information. Our paper shows the converse: When investors can choose what to learn about, the value of diversification declines.

home investor to acquire information that magnifies his comparative advantage in home assets. If home investors undo their information asymmetry by learning about foreign assets, they earn no excess returns. When information indicates that an asset's payoff will be high, all investors know about it and bid up the price. If an investor instead learns more about some assets and less about others than what the average investor learns about, and then takes a large position in the assets he knows more about, he will earn higher profits. When information about an asset he is better informed about indicates that the payoff will be high, the price need not be high, allowing him to profit. When choosing what information to learn about, investors' goal is to make their information sets as different as possible from that of the average investor. The most efficient way to achieve that goal is for home investors to take the home assets they start out knowing relatively more about and specialize in learning even more about them. The main result in the first half of the paper is that information immobility persists not because investors can't learn what locals know, nor because it is too expensive, but because they don't choose to; specializing in what they already know is a more profitable strategy. Having shown that sustaining information asymmetry is possible, the second half of the paper compares the testable predictions of the model to the data.

The model's key mechanism is the interaction between information and investment choice. To illustrate its importance, section 2 shuts down this interaction by forcing investors to take their portfolios as given, when they choose what to learn. These investors minimize investment risk by learning about assets that they are most uncertain about. With sufficient capacity, learning undoes all initial information advantage, and therefore all home bias. Thus, this model embodies the logic that the asymmetric information criticisms are founded on.

Section 3 shows that when investors have rational expectations about their future optimal portfolio choices, this logic is reversed. While acquiring information that others do not know increases expected portfolio returns, it does not imply that home investors take a *long* position in home assets, only that they take a *large* position. Home bias arises because home assets offer risk-adjusted expected excess returns to informed home investors. Information about the home asset reduces the risk that the asset poses without changing its return, hence the high risk-adjusted returns. Why does information reduce risk? An asset's payoff may be very volatile, high one period and low the next. But if an investor has information that tells him when the payoff is high and when it is low, the asset is not very risky to that investor. While foreign assets offer lower risk-adjusted returns to home investors, they are still held for diversification purposes. The optimal portfolio is tilted toward home assets.

Considering how learning affects portfolio risk offers an alternative way of understanding why investors with an initial information advantage in home assets choose to learn more about home assets. Because of the excess risk-adjusted returns, a home investor with a small information advantage initially expects to hold slightly more home assets than a foreign investor would. This small initial difference is amplified because information has increasing returns in the value of the asset it pertains to: as the investor decides to hold more of the asset, it becomes more valuable to learn about. So, the investor chooses to learn more and hold more of the asset, until all his capacity to learn is exhausted on his home asset.

A variety of evidence supports the model's predictions. Section 4 connects the theory to facts about analyst forecasts, portfolio patterns, excess portfolio returns, cross-sectional asset prices, as well as evidence thought to be incompatible with an information-based home bias explanation. In particular, the theory offers a unified explanation of home bias and more recent findings of local bias. While we cannot claim for any one of these facts that no other theory could possibly explain the same relationship, taken together, they constitute a large body of evidence that is consistent with one parsimonious theory. A numerical example shows that learning can magnify the home bias considerably. When all home investors get a small initial advantage in all home assets, the home bias is between 5 and 46%, depending on the magnitude of investors' learning capacity. When each home investor gets an initial information advantage that is concentrated in one local asset, the home bias is amplified. It rises as high as the 76% home bias in U.S. portfolio data, for a level of capacity that is consistent with observed excess returns on local assets. Finally, we derive new testable hypotheses from the model to guide future empirical work.

Information advantages have been used to explain exchange rate fluctuations (Evans and Lyons (2004), Bacchetta and van Wincoop (2006)), the international consumption correlation puzzle (Coval 2000), international equity flows (Brennan and Cao 1997), a bias towards investing in local stocks (Coval and Moskowitz 2001), and the own-company stock puzzle (Boyle, Uppal and Wang 2003). Information asymmetry also sustains other home bias explanations, such as ambiguity aversion (Uppal and Wang 2003). All of these explanations are bolstered by our finding that information advantages are not only sustainable when information is mobile, but that asymmetry can be amplified when investors can choose what to learn.

1 A Model of Learning and Investing

Using tools from information theory, we construct an equilibrium framework to consider learning and investment choices jointly. This model uses the one-investor partial-equilibrium problem of Van Nieuwerburgh and Veldkamp (2006) to build a heterogeneous-agent, two-country equilibrium model with a continuum of investors in each country. Modeling equilibrium interactions delivers important new insights. In particular, it allows us to show that investors want to make their information different from what other investors know. This is the force that delivers asymmetric information and home bias. Furthermore, the initial heterogeneity in investors across countries delivers predictions about which investors learn what and allows this model to connect with portfolio data.

This is a static model which we break up into 3 periods. In period 1, investors choose the distribution from which to draw signals about the payoff of the assets, subject to a constraint on the total informativeness of their signals. In period 2, each investor observes signals from the chosen distribution and makes his investment. Prices are set such that the market clears. In period 3, he receives the asset payoffs and consumes.

Preferences Investors, with absolute risk aversion parameter ρ, facing an $N \times 1$ vector of unknown asset payoffs f a risk-free rate r and asset prices p, maximize their mean-variance utility:

$$U = -E\left[-\rho q'(f - rp) + \frac{\rho^2}{2} q'\widehat{\Sigma} q\right]. \tag{1}$$

where q is the $N \times 1$ vector of quantities of each asset the investor decides to hold and $\widehat{\Sigma}$ is the uncertainty about payoffs that investors face after they learn.[3] When portfolios are chosen in period 2, the expectation E is conditional on the realization of the signals the investor has chosen to see. When signals are chosen at time 1, the investor does not know what the realizations of these

[3] A technical appendix, posted on the authors' websites, discusses the foundations for this utility formulation in detail. The results do not depend on the existence of a risk-free asset. Suppose investors can consume c_1 at the investment date and c_2 when asset payoffs are realized. If preferences are defined over $rc_1 + c_2$, where r is the rate of time preference, the solution will be identical. The earlier consumption choice takes the place of the riskless investment choice.

signals will be. Therefore, in period 1, the investor has the same objective, except that expectation E conditions only on information in prior beliefs. This utility function comes from an exponential form of utility over terminal wealth. Terminal wealth equals initial wealth W_0, plus the profit earned from portfolio investments:

$$W = rW_0 + q'(f - pr) \tag{2}$$

Initial information Two countries, home and foreign, have an equal-sized continuum of investors, whose preferences are identical. Investors are endowed with prior beliefs about a vector of asset payoffs f. Each investor's prior belief is an unbiased, independent draw from a normal distribution, whose variance depends on where the investor resides. Home prior beliefs are $\mu \sim N(f, \Sigma)$. Foreign prior beliefs are distributed $\mu^\star \sim N(f, \Sigma^\star)$. Home investors have lower-variance prior beliefs for home assets and foreign investors have lower-variance beliefs for foreign assets. One interpretation is that each investor gets a free signal about each asset in his home country. We will call this initial difference in variances a group's information advantage.

Information acquisition For intuition, think of each investor as an econometrician who knows the true payoff's mean and variance of asset payoffs f. The only unknown is the realization of those payoffs, which is what the investor can learn about. He can acquire additional data to form a more accurate payoff estimate $\hat{\mu}$. The investor chooses what assets to collect data on, subject to a constraint on the total amount of data. Collecting more data on one asset reduces the standard error of his estimate for that asset's payoff. The posterior variance is that standard error, squared.

At time 2, each investor will observe an $N \times 1$ vector of signals η about the vector of asset payoffs f. At time 1, investors choose what kind of signals to acquire. They don't choose whether signals will contain good or bad news. Rather, they choose signals that will contain more precise information about some assets than others. In other words, they choose a variance Σ_η such that $\eta \sim N(f, \Sigma_\eta)$. Each investor's signal is independent of the signals drawn by other investors.

When payoffs co-vary, obtaining a signal about one asset's payoff is informative about other payoffs. To describe what a signal is about, it is useful to decompose asset payoff risk into orthogonal risk factors and the risk of each factor. This decomposition breaks the prior variance-covariance matrix Σ up into a diagonal eigenvalue matrix Λ, and an eigenvector matrix Γ: $\Sigma = \Gamma\Lambda\Gamma'$. The

Λ_i's are the variances of each risk factor i. The ith column of Γ (denoted Γ_i) gives the loadings of each asset on the ith risk factor. To make aggregation tractable, we assume that home and foreign prior variances Σ and $\Sigma^\star$ have the same eigenvectors, but different eigenvalues. In other words, home and foreign investors use their capacity to reduce different initial levels of uncertainty about the same set of risks. This assumption implies that investors observe signals ($\Gamma'\eta$) about risk factor payoffs ($\Gamma' f$). Learning about risk factors (principal components analysis) has long been used in financial research and among practitioners. It approximates risk categories investors might study: country risk, business cycle risk, industry, regional, and firm-specific risk. Nothing prevents investors from learning about many risk factors. The only thing this rules out is signals with correlated information about independent risks.

Choosing how much to learn about each risk factor is equivalent to choosing the variance of each entry of the N-dimensional signal vector $\Gamma'\eta$. Since the signal is unbiased, its mean is $\Gamma' f$. The variance of a principal component is its eigenvalue. So, reducing uncertainty about the ith risk factor means choosing a smaller ith eigenvalue of the signal variance-covariance matrix Σ_η. Signals about the payoffs of all assets that load on risk factor i become more accurate. With Bayesian updating, each Σ_η results in a unique posterior variance matrix that measures the investor's uncertainty about asset payoffs, after incorporating what he learned. Since the mapping between signal choices and posteriors is unique, information choice is the same as choosing posterior variance, without loss of generality. Since sums, products and inverses of prior and signal variance matrices have eigenvectors Γ, posterior beliefs will as well. Denoting posterior beliefs with a hat, $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$, where Γ is given and the diagonal eigenvalue matrix $\hat{\Lambda}$ is the choice variable. The decrease in risk factor i's posterior variance ($\Lambda_i - \hat{\Lambda}_i$) measures the decrease in uncertainty achieved through learning.

There are 2 constraints governing how the investor can choose his signals about risk factors. The first is the *capacity constraint*; it limits the quantity of information investors can observe. Grossman and Stiglitz (1980) used the ratio of variances of prior and posterior beliefs to measure the 'quality of information' about one risky asset. We generalize the metric to a multi-signal setting by bounding the ratio of the generalized prior variance to the generalized posterior variance, $|\widehat{\Sigma}| \geq \frac{1}{K}|\Sigma|$, where generalized variance is a term that refers to the determinant of the variance-covariance matrix. Capacity $K \geq 1$ measures how much an investor can decrease the uncertainty he faces. For now, K is the same for all investors. Since determinants are a product of eigenvalues, the capacity constraint is

$$\prod_i \hat{\Lambda}_i \geq \frac{1}{K} \prod_i \Lambda_i. \tag{3}$$

The second constraint is the *no negative learning constraint*: the investor cannot choose to increase uncertainty (forget information) about some risks to free up more capacity to decrease uncertainty about other risks. We rule this out by requiring the variance-covariance matrix of the signal vector $\Sigma_\eta = \Gamma\Lambda_\eta\Gamma'$, to be positive semi-definite. Since a matrix is positive semi-definite when all its eigenvalues are positive, the constraint is:

$$\Lambda_{\eta i} \geq 0 \quad \forall\, i. \tag{4}$$

Comments on the learning technology The structure we put on the learning problem keeps it as simple as possible. But many of these assumptions can be relaxed. First, our results do not hinge on the assumption that investors learn about principal components of asset payoffs. Investors specialize in what they know well, for any arbitrary risk factor structure. Second, our framework can incorporate heterogeneous capacity (see section 4.3). Third, allowing agents to choose how much capacity to acquire does not change the results. Any cost function increasing in K has an equivalent capacity endowment that produces identical portfolio outcomes. Finally, a learning technology with diminishing returns and un-learnable risk will moderate, but not overturn, our results. Instead of specializing in one risk, investors may learn about a limited set of risks. But it does not change the conclusion that investors prefer to learn about what they already have an advantage in.[4]

It is not true that every capacity constraint preserves specialization. We use this one because it is a common distance measure in econometrics (a log likelihood ratio) and in statistics (a Kullback-Liebler distance); it is a bound on entropy reduction, an information measure with a long history in information theory (Shannon 1948); it can be interpreted as a technology for reducing measurement error (Hansen and Sargent 2001); it is a measure of information complexity (Cover and Thomas 1991); it has been used to forecast foreign exchange returns (Glodjo and Harvey 1995), and it has been used to describe limited information processing ability in economic settings by (Sims 2003).[5]

[4] A proof of the first, third and final claims can be found in the technical appendix, posted on the authors' websites.

[5] This learning technology is also used in models of rational inattention. However, that work has focused on time-series phenomena in representative investor models such as delayed response to shocks, inertia, time to digest, and consumption smoothing. See e.g. Sims (2003) and Moscarini (2004). Instead, we focus on the strategic interactions

Although we do not prove this is the correct learning technology, our strategy is to work out its predictions for international investment choices and ask whether they are consistent with the data.

Updating beliefs When investors' portfolios are fixed (section 2), what investors learn does not affect the market price. But when asset demand responds to observed information (section 3), the market price is an additional noisy signal of this aggregated information. Using their prior beliefs, their chosen signals, and information contained in prices, investors form posterior beliefs about asset payoffs, using Bayes' law.

The information in prices depends on portfolio choices. Appendix A.3 shows that prices are linear functions of the true asset payoffs such that $(rp - A) \sim N(f, \Sigma_p)$, for some constant A.

An investor j's posterior belief about the asset payoff f, conditional on a prior belief μ^j, signal $\eta^j \sim N(f, \Sigma_\eta^j)$, and prices, is formed using Bayesian updating. The posterior mean is a weighted average of the prior, the signal and price information, while the posterior variance is a harmonic mean of the prior, signal, and price variances:

$$\hat{\mu}^j \equiv E[f|\mu^j, \eta^j, p] = \left((\Sigma^j)^{-1} + (\Sigma_\eta^j)^{-1} + \Sigma_p^{-1}\right)^{-1} \left((\Sigma^j)^{-1}\mu^j + (\Sigma_\eta^j)^{-1}\eta^j + \Sigma_p^{-1}(rp - A)\right) \tag{5}$$

$$\hat{\Sigma}^j \equiv V[f|\mu^j, \eta^j, p] = \left((\Sigma^j)^{-1} + (\Sigma_\eta^j)^{-1} + \Sigma_p^{-1}\right)^{-1}. \tag{6}$$

We emphasize that acquiring information $((\Sigma_\eta^j)^{-1} > 0)$ always reduces posterior variance. This might appear puzzling because in an econometric setting, new data can make us revise our variance estimates upward. The difference is that there is no estimation of variance in our problem. The true variance of f is known to all investors. Rather, $\widehat{\Sigma}$ is the variance of the estimate of f. It is a measure of uncertainty, not of volatility. Under Bayes' law with normal random variables, more information always reduces uncertainty.

Market clearing Asset prices p are determined by market clearing. The per-capita supply of the risky asset is $\bar{x} + x$, a positive constant ($\bar{x} > 0$) plus a random ($n \times 1$) vector with known mean and variance, and zero covariance across assets: $x \sim N(0, \sigma_x^2 I)$. The reason for having a risky asset supply is to create some noise in the price level that prevents investors from being able to perfectly infer the private information of others. Without this noise, no information would be private, and

and heterogeneity of individuals' learning choices.

no incentive to learn would exist. We interpret this extra source of randomness in prices as due to liquidity or life-cycle needs of traders. The market clears if investors' portfolios q^j sum to the asset supply: $\int_0^1 q^j dj = \bar{x} + x$.

Definition of Equilibrium An equilibrium is a set of asset demands, asset prices and information choices, such that three conditions are satisfied. First, given prior information about asset payoffs $f \sim N(\mu, \Sigma)$, each investor's information choice $\hat{\Lambda}$ and portfolio choice q maximize (1), subject to capacity (3), no-negative-learning (4) and budget constraints (2). Second, asset prices are set such that the asset market clears. Third, beliefs are updated, using Bayes' law: (5) and (6) and expectations are rational: Period-1 beliefs about the portfolio q are consistent with the true distribution of the optimal q.

2 Why Might Asymmetric Information Disappear?

Returns to specialization come from the interaction of the investment choice and the learning choice. To highlight the importance of this interaction, we first explore a model where it is shut down. The only difference with the model in section 3 is that investors do not account for the fact that what they learn will influence the portfolio they hold. They choose what to learn, in order to minimize the risk of a portfolio that they take as given. In this setting, investors learn exclusively about the most uncertain assets until either they run out of capacity, or are equally uncertain about all assets. Learning undoes information asymmetry and reduces or eliminates home bias. As Karen Lewis (1999) put it, "Greater uncertainty about foreign returns may induce the investor to pay more attention to the data and allocate more of his wealth to foreign equities." This section explains the basis for her criticism. The next section exposes its logical flaw.

A Model without Increasing Returns to Information In order to shut down the investment-learning interaction, suppose the investor takes the vector of asset holdings q as given, when choosing what to learn. Define the amount of risk factor i that an investor holds in his portfolio as $\tilde{q}_i = \Gamma_i' q$. Then the objective (1) collapses to choosing $\hat{\Lambda}_i$'s to minimize $\sum_i \tilde{q}_i^2 \hat{\Lambda}_i$, subject to the capacity constraint (3) and the no-forgetting constraint $\Lambda_i - \hat{\Lambda}_i \geq 0 \ \forall i$. The following result shows that learning undoes initial information asymmetry.[6]

Proposition 1. *If an investor has an informational advantage in one risk factor $\Lambda_i < \Lambda_j \ \forall j$, then with sufficient information capacity $K \geq K^*$, the investor will choose the same posterior variance that he would choose if his advantage was in any other risk factor: $\Lambda_k < \Lambda_j \ \forall j$ for some $k \neq i$.*

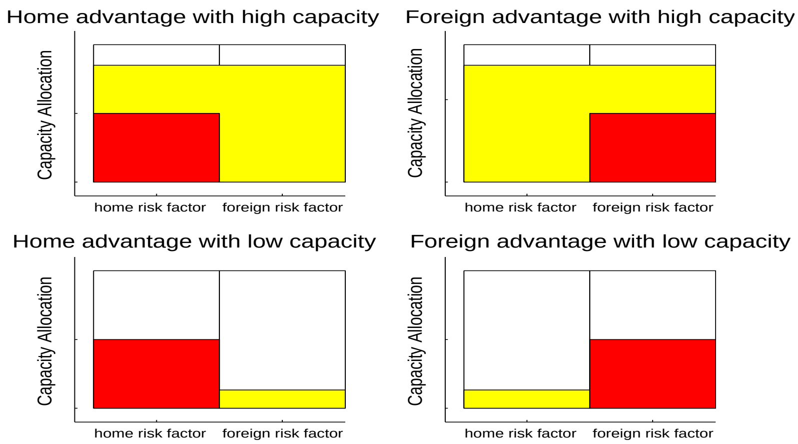

Figure 1: Allocation of information capacity for a low and high-capacity investor.
The lightly shaded area represents the amount of capacity allocated to the factor. The dark area represents the size of the information advantage. The unfilled part of each bin represents the posterior variance of the risk factor $\hat{\Lambda}_i$. With high capacity, adding the dark block to either bin would result in the 'water level' $\hat{\Lambda}^{-1}$ being the same for both risk factors. This is the case where initial information advantages are undone by learning.

The top two panels of figure 1 illustrate this corollary. The brick and water picture is a metaphor for how information capacity (the water) is diverted to other risks when an investors have an initial information advantage (the brick). There is a home and foreign risk factor (h, f); the two bins are equally deep because both risks are equally valuable to learn about $(\tilde{q}^h = \tilde{q}^f)$. Giving an investor a home (foreign) information advantage is like placing a brick in the left (right) side of the box. When capacity is high, a brick placed on either side will raise the water level on both sides equally. Having an initial advantage in home or foreign risk will result in the same the same posterior variances for both assets. Learning choices compensate for initial information advantage in such a way as to render the nature of the initial advantage irrelevant. Any home bias that might result from the information advantage disappears when investors can learn.

The bottom panels of figure 1 illustrate low-capacity allocations. The investor would like the water level (his posterior precision) to be the same in both bins, but there is insufficient water (capacity). The no-forgetting constraint prevents him from breaking up the brick to level the

[6]The proofs of this and all subsequent propositions are in appendix A.

water. He cannot equalize home and foreign uncertainty. The constrained optimum is to devote all capacity to the most uncertain risk. For a home investor with an initial advantage in the home risk factor, this means she should use all capacity to learn about the foreign risk factor.

Initial information advantages could persist if capacity were low relative to the initial advantage (as in the bottom panels of figure 1). However, if this explanation were true, then individuals would never choose to learn about home assets; they would devote what little information capacity they had entirely to learning about foreign assets. This implication is inconsistent with the multi-billion-dollar industry that analyzes U.S. stocks, produces reports on the U.S. economy, manages portfolios of U.S. assets, and then sells their products to American investors.

A second mechanism that might preserve a home information advantage is a higher cost of processing foreign information. While foreign information is likely harder to learn, this cost difference must be large to account for the magnitude of the home bias. Since there is no theory to predict information costs and they are not observable, it is desirable for a theory not to rely on the magnitude of the cost difference. Instead, the model in the next section requires an arbitrarily small initial information advantage, possibly generated by a small cost difference, to endogenously create a large home bias.

3 Main Results

The previous section illustrated how information asymmetry could disappear. This section analyzes a model where small asymmetries in investors' information not only persist, but are magnified. The only change in the setup is that investors do not take their asset demand, or the asset demand of other investors, to be fixed. Instead, we apply rational expectations: every investor takes into account that every portfolio in the market depends on what each investor learns. We conclude that home investors can learn foreign information, but choose not to. They make more profit from specializing in what they already know.

The Period-2 Portfolio Problem We solve the model using backward induction, starting with the optimal portfolio decision, taking information choices as given. Given posterior mean $\hat{\mu}^j$ and variance $\widehat{\Sigma}^j$ of asset payoffs, the portfolio for investor j, from either country, is

$$q^j = \frac{1}{\rho}(\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr). \tag{7}$$

Aggregating these asset demand across investors and imposing the market clearing condition delivers a solution for the equilibrium asset price level that is linear in the asset payoff f and the unexpected component of asset supply x: $p = \frac{1}{r}(A + f + Cx)$. Appendix A.3 derives formulas for A and C.

The Optimal Learning Problem In period 1, the investor chooses information to maximize expected utility. In order to impose rational expectations, we substitute the equilibrium asset demand (7), into expected utility (1). Combining terms yields

$$U = E\left[\frac{1}{2}(\hat{\mu}^j - pr)'(\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr)|\mu, \Sigma\right]. \tag{8}$$

At time 1, $(\hat{\mu}^j - pr)$ is a normal variable, with mean $(-A)$ and variance $\Sigma_p - \widehat{\Sigma}^j$.[7] Thus, expected utility is the mean of a chi-square. Using the fact that the choice variable $\hat{\Lambda}$ is a diagonal matrix, that $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$, the formula for A (18), and the formula for the mean of a chi-square, we can rewrite the period-1 objective as:

$$\max_{\hat{\Lambda}^j} \sum_i \left(\Lambda_{pi} + (\rho\Gamma_i'\bar{x}\hat{\Lambda}_i^a)^2\right)(\hat{\Lambda}_i^j)^{-1} \quad s.t.\ (3) \text{ and } (4) \tag{9}$$

where Λ_{pi} is the ith eigenvalue of Σ_p, and $\hat{\Lambda}_i^a = (\int_j (\hat{\Lambda}^j)^{-1})^{-1}$ is the posterior variance of risk factor i of a hypothetical investor whose posterior belief precision is the average of all investors' precisions.

The key feature of the learning problem (9) is its convexity in the posterior variance $(\hat{\Lambda}^j)$. In a 2-risk factor setting, the objective is $U = L_1/\hat{\Lambda}_1 + L_2/\hat{\Lambda}_2$, for positive scalars L_1, L_2. Thus, an indifference curve is $\hat{\Lambda}_2 = L_2\hat{\Lambda}_1/(U\hat{\Lambda}_1 - L_1)$, which asymptotes to ∞ at $\hat{\Lambda}_1 = L_1/U > 0$. The capacity constraint is $\hat{\Lambda}_2 = K/\hat{\Lambda}_1$, which asymptotes to ∞ at $\hat{\Lambda}_1 = 0$. Because the indifference curve is always crossing the capacity constraint from below, the solution is always a corner solution. Figure 2 plots the indifference curve (for $L_1 = L_2$), the capacity constraint, and the no-negative

[7] To derive this variance, note that $\text{var}(\hat{\mu}|\mu) = \Sigma - \widehat{\Sigma}$, that $\text{var}(pr|\mu) = \Sigma + \Sigma_p$, and that $\text{cov}(\hat{\mu}, pr) = \Sigma$.

learning bounds for our model (left panel) and the exogenous-portfolio model in section 2 (right panel).

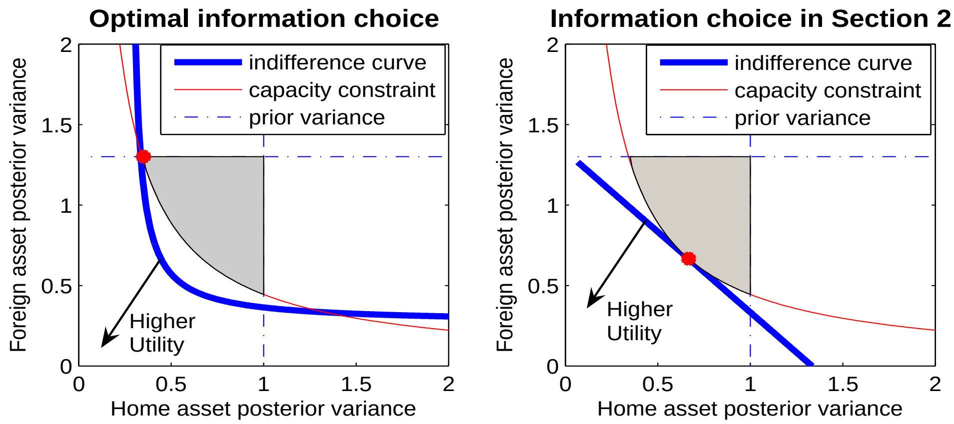

Figure 2: Objective and constraints in the optimal learning problem with 2 risk factors.

Utility increases as the indifference curve (dark line) moves toward the origin (variance falls). All feasible learning choices must lie on or above the capacity constraint (lighter line). The no-negative learning constraint further prohibits posterior variances from exceeding prior variances (dashed lines). The set of learning choices that satisfy both constraints is the shaded set. Whenever foreign prior variance is higher than home prior variance, the solution in our model (the large dot in the left panel) is to devote all capacity to reducing home asset risk. In the section 2 model (right panel), the objective is linear and the optimum is to reduce variance on home and foreign assets. The right panel shows why shutting down the information-portfolio interaction reverses our main conclusion.

Proposition 2. ***Optimal Information Acquisition.*** *Each investor j uses all capacity to learn about one linear combination of asset payoffs. The linear combination is the payoff of risk factor i $f'\Gamma_i$ associated with the highest value of the* ***learning index****: $\frac{\hat{\Lambda}_i^a}{\Lambda_i^j}\rho^2(\Gamma_i'\bar{x})^2\hat{\Lambda}_i^a + \frac{\Lambda_{pi}}{\Lambda_i^j}$.*

Three features make a particular risk factor i desirable to learn about. First, since information has increasing returns, the investor gains more from learning about a risk that is abundant (high $(\Gamma_i'\bar{x})^2$). Second, the investor should learn about a risk factor that the average investor is uncertain about (high $\hat{\Lambda}_i^a$). These risks have prices that reveal less information (high Λ_{pi}), and higher returns: $\Gamma_i'E[f - pr] = \rho\hat{\Lambda}_i^a\Gamma_i'\bar{x}$. (See appendix A.3 for definitions of $\hat{\Lambda}^a$, Λ_p and expected returns.) Third, and most importantly for the point of the paper, the investor should learn about risk factors that he has an initial advantage in, relative to the average investor (high $\hat{\Lambda}_i^a/\Lambda_i$). Since these are the assets he will expect to hold more of, these are more valuable to learn about.

The feedback effects of learning and investing can be seen in the learning index. The amount of a risk factor that an investor j expects to hold, based on his prior information, is the factor's expected return, divided by its prior variance: $(\Lambda_i^j)^{-1}\rho\hat{\Lambda}_i^a\Gamma_i'\bar{x}$. This expected portfolio holding shows up in the learning index formula, indicating that a higher expected portfolio share increases the value of learning about the risk. Expecting to learn more about the risk lowers the posterior variance $\hat{\Lambda}_i^j$. Re-computing the expected portfolio with variance $\hat{\Lambda}$, instead of Λ, further increases factor i's portfolio share, and feeds back to increase i's learning index. This interaction between the learning choice and the portfolio choice, an endogenous feature of the model, generates increasing returns to specialization.

Strategic Substitutability Because other investors' learning lowers the $\hat{\Lambda}_i^a$ and Λ_{pi} for the risks they learn about, each investor prefers to learn about risks that others do not learn. This Nash equilibrium could be reached by an iterative choice process. The first investor begins by learning about the risk with the highest learning index. If another risk factor l has a learning index is not far below that of i, then the fall in $\hat{\Lambda}_i^a$, brought on by some investors learning about i will cause other investors to prefer learning about l. If all home investors are ex-ante identical, they will be indifferent between learning about any of the risks that any home investor learns about. Foreign investors will also be indifferent between any of the risks that foreigners learn about. Although investors may be indifferent between specializing in any one of many risk factors, the aggregate allocation of capacity is unique. The number of home and foreign risk factors learned about in each country depends on country-wide capacity. Despite the fact that many risk factors are potentially being learned about in equilibrium, each investor learns about only one factor.

Learning and Information Asymmetry The effect of an initial information advantage on a learning is similar to the effect of a comparative advantage on trade. Home investors always have a higher learning index than foreigners do for home risks, and vice-versa for foreign risks. If home risks are particularly valuable to learn about, for example because those risks are large (high $\Gamma_i'\bar{x}$), some foreigners may choose to learn about them. But, if home risks are valuable to learn about, all home investors will specialize in them. Likewise, if some home investors learn about foreign risks, then all foreigners must be specializing in foreign risks as well. The one pattern the model rules out is that home investors learn about foreign risk and foreigners learn about home risk. This is analogous to the principle of comparative advantage: If country A has an advantage in producing apples and country B an advantage in bananas, the one production pattern that is not possible is that country A produces bananas and B apples. Investors never make up for their initial information asymmetry by each learning about the others' advantage. Instead, posterior beliefs diverge, relative to priors; information asymmetry is amplified.

To describe this result formally, we need some new notation. Let $\Lambda_h, \Lambda_f, \hat{\Lambda}_h$ and $\hat{\Lambda}_f$ be $N/2$-by-$N/2$ diagonal matrices that lie on the diagonal quadrants of the prior and posterior belief matrices: $\Lambda = [\Lambda_h 0; 0 \Lambda_f]$ and $\hat{\Lambda} = [\hat{\Lambda}_h 0; 0 \hat{\Lambda}_f]$. And, let the * superscript on each of these matrices denotes foreign belief counterparts. Then, for example, Λ_f represents home investors' prior uncertainty about foreign risk factors and $\hat{\Lambda}^\star_h$ represents foreigners' posterior uncertainty about home risks.

Proposition 3. ***Learning Amplifies Information Asymmetry.*** *Learning will amplify initial differences in prior beliefs for every pair of home and foreign investors:* $\frac{|\hat{\Lambda}^\star_h|}{|\hat{\Lambda}_h|} \geq \frac{|\Lambda^\star_h|}{|\Lambda_h|}$ *and* $\frac{|\hat{\Lambda}_f|}{|\hat{\Lambda}^\star_f|} \geq \frac{|\Lambda_f|}{|\Lambda^\star_f|}$.

A special case of this result arises when home and foreign countries are perfectly symmetric: They have an equal number of risk factors of equal size, with equal payoff variances. In this case, home investors will learn exclusively about home risks and foreign investors will learn exclusively about foreign risks. This result follows directly from the learning index in proposition 2. An investor with no information advantage would have identical learning indices for home and foreign risks. Thus, he would be indifferent between learning about home and foreign risks. Since investors with no information advantage are indifferent, *any initial advantage* in home risk i (lower Λ^j_i) breaks that indifference, tilts preferences toward learning more about home risk and amplifies the initial advantage.

At the other extreme, with very asymmetric markets, amplification disappears. If the home market is much smaller than foreign, all investors could learn about foreign risk factors. The ratio of home and foreign investors' posterior precisions will then be the same as the ratio of their prior precisions. The initial advantage will just be preserved.

Home Bias in Investors' Portfolios To explore the effect of learning on home bias, we compare our model's predictions to two benchmark portfolios. The first portfolio is one with no information advantage and no capacity to learn. Home investors and foreign investors have identical beliefs and hold identical portfolios, which depend on the random asset supply. The expected portfolio is the per capita expected supply: $E[q^{no\ adv}] = \bar{x}$.

A second natural benchmark portfolio is one where investors have initial information advantages, but no capacity to acquire signals and do not learn through prices ($K = 1$). For example, this is the kind of information advantage that Ahearne et al. (2004) capture when they estimate the home bias that uncertainty about foreign accounting standards could generate. $E[q^{no\ learn}] = \Gamma\Lambda^{-1}\Lambda^{a}\Gamma'\bar{x}$, where Λ^a is the average investor's *prior* variance.

Specialization in learning does not imply that the investors hold exclusively home assets. They still exploit gains from diversification. Each investor's portfolio takes the world market portfolio ($\bar{x}$) and tilts it towards the assets i that he knows more about than the average investor (high $\hat{\Lambda}_i^{-1}\hat{\Lambda}_i^a$). The optimal expected portfolio with learning is

$$E[q] = \Gamma\hat{\Lambda}^{-1}\hat{\Lambda}^{a}\Gamma'\bar{x} \tag{10}$$

Learning has two effects on an investors' portfolio. First, it magnifies the asset position and second, it tilts the portfolio towards the assets learned about. The first effect can be seen in (10): Learning increases the precision of beliefs $\hat{\Lambda}^{-1} > \Lambda^{-1}$. Lower risk in factor i makes investors want to take larger positions in i, positive or negative. But why should the position in home assets be a large long position, rather than a large short one? The second effect is an equilibrium effect. The return on an asset compensates the average investor for the amount of risk he bears $\hat{\Lambda}_i^a$. The fact that foreign investors are investing in home assets without knowing much about them (typically as part of a diversified portfolio), raises $\hat{\Lambda}^a$ and thus the asset's return. Home investors are being compensated for more risk than they bear ($\hat{\Lambda}_i^a > \hat{\Lambda}_i^j$). In other words, the home assets deliver high risk-adjusted returns. High returns make a long position optimal, on average. Both the magnitude and the general equilibrium effect increase home bias.[8]

The final proposition compares home bias in the optimal portfolio (10) and in the benchmark portfolios. Let Γ_h be a sum of the eigenvectors in Γ which correspond to the home risk factors. Then $\Gamma_h'q$ quantifies how much total home risk an investor is holding in their portfolio.

[8] It is possible that a highly negative signal realization on a home asset would make home investors who are informed short that asset. Short selling is unlikely to occur on a large scale in general equilibrium. The dramatic fall in prices from widespread shorting would signal the bad news to foreign investors, making them unwilling to take the corresponding large long positions. Low prices would also make home investors more willing to hold home assets, despite their low payoffs.

Proposition 4. ***Learning Increases Home Bias.*** *The average home investor's portfolio contains at least as much of assets that load on home risk when he can learn ($K > 1$), than when he cannot ($K = 1$):* $\Gamma_h' E[q] \geq \Gamma_h' E[q^{no\ learn}] > \Gamma_h' E[q^{no\ adv}]$.

4 Bringing the Theory to Data

There are a number of recent papers that present alternative explanations for home bias. Some of these explanations are behavioral: Huberman (2001) explores familiarity, Cohen (2004) explores loyalty, Morse and Shive (2003) propose patriotism, while Graham, Harvey and Huang (2006) investigate overconfidence. Other argue, like this paper does, that home bias is optimal: Cole, Mailath and Postlewaite (2001) and DeMarzo, Kaniel and Kremer (2004) claim that investors have preference-based or market-price-based incentives to hold portfolios similar to their neighbors'. At the same time, there has been an active literature that attempts to distinguish between the various theories by documenting facts related to the home bias. Our theory offers a parsimonious explanation for many of these facts. Rather than adding new facts, this section taps in to the existing empirical literature and connects the theory to the evidence, qualitatively and quantitatively (sections 4.1 and 4.2). It also reconciles existing facts that appear to be at odds with an information explanation (section 4.3) and offers new predictions that can guide future empirical work (section 4.4).

4.1 Facts That Support Model Predictions

Direct Evidence of Information Asymmetry Bae, Stulz and Tan (2005) measure information asymmetry and link it to home bias. They show that home analysts in 32 countries make more precise earnings forecasts for home stocks than foreign analysts do. On average, the increase in precision is 8%. Furthermore, the size of the home analyst advantage is related to home bias. When local analysts' forecasts are more precise relative to foreigners' forecasts (more information asymmetry), foreign investors hold less of that country's assets.

Guiso and Jappelli (2006) examine survey data on the time customers of a leading Italian bank spend acquiring financial information. Those who spend more time on information collection hold portfolios that are less diversified and earn significantly higher returns.

Local Bias Home bias is not just a country-level effect. Investors also favor local assets, headquartered near their home, over firms in the same country located further away (Coval and Moskowitz 2001). A unified explanation for home and local bias is something that many theories cannot provide. Their coexistence makes an information-based explanation appealing. Malloy (2005) offers direct evidence that local analysts do in fact have information advantages. He shows that local analysts' forecasts better predict stock returns and that they earn abnormal returns on their local assets. By giving investors slightly more precise initial information about local assets, this model can explain the local bias.

Suppose that home investors each had an advantage in only one home risk factor, the one most concentrated in their region's asset. An investor j from region m draws an independent prior belief $\mu^j \sim N(f, \Sigma_m)$, where $\Sigma_m = \Gamma\Lambda_m\Gamma$, and Λ_m has a mth diagonal entry that was lower than the mth diagonal in the beliefs of any other region. In this model, local investors have an incentive to learn more about their local assets, because of their initial information advantage (proposition 2). Local advantages also *amplify the effects* of home advantages: When fewer investors share an advantage in the same local risk, locals have a larger advantage relative to the average investor (higher $\hat{\Lambda}^a_m/\Lambda^j_m$). A more specialized advantage magnifies the optimal portfolio bias ($E[\Gamma'_m q] = \hat{\Lambda}^a_m/\Lambda^j_m(\Gamma'_m\bar{x})$). Because returns to specialization increase when information advantages are more concentrated, investors diversify less. We illustrate this amplification effect in section 4.2.

Industry Bias One source of prior information advantages could be one's industry. If so, investors should reinforce that information asymmetry by learning more about that industry and investing more in it. Massa and Simonov (2006) confirm this prediction. They show that Swedish investors buy assets closely related to their non-financial income. Two facts suggest that this behavior is not a bias, but is information-driven. First, when an investor changes industries, the industry concentration of his portfolio declines. This is consistent because information takes time to accumulate. Second, "familiarity-based" portfolios yield higher returns than diversified ones.

Another source of prior information is one's classmates. Cohen, Frazzini and Malloy (2007) find that fund managers over-invest in firms run by their former classmates and make excess returns on those investments. This is consistent with an initial information advantage acquired in school.

Under-diversified Foreign Investment One feature of portfolio data that is difficult to explain is the concentration within the foreign component of home investors' portfolios. The part of a portfolio invested in any given foreign country should therefore be diversified. Kang and Stulz (1997) show that this is not the case. Using data on foreign investors in Japan, they show that foreigners' portfolios of Japanese assets overweight large firms and assets whose returns correlate highly with aggregate risk.

This pattern is consistent with our model. Suppose than an American investor chooses to learn about and invest in Japanese assets. Holding equal the average uncertainty ($\hat{\Lambda}^a$), noise in prices (Λ_p) and American prior uncertainty (Λ) about each Japanese risk, the most valuable risk to reduce is the one with the largest quantity (highest $\Gamma_i \bar{x}$ in proposition 2). In other words, the American should learn about the largest risk factors, aggregate macroeconomic risk and the risks associated with the largest firms. Since investors, on average, hold more of the assets they've learned about, the model predicts that Americans who hold Japanese assets will not diversify their Japanese holdings. Instead, they will overweight large, high-beta firms.

Portfolio Out-performance If transaction costs or behavioral biases are responsible for under-diversification, then concentrated portfolios should deliver no additional profit. In contrast, if investors in our model concentrate their portfolios, it is because they have informational advantages. Their concentrated portfolios should out-perform diversified ones.[9]

There is empirical evidence for such out-performance. Ivkovic, Sialm and Weisbenner (2007) find that concentrated investors outperform diversified ones by as much as 3% per year. Out-performance is even higher for investments in local stocks, where natural informational asymmetries are most likely to be present (see also Coval and Moskowitz (2001), Massa and Simonov (2006) and Ivkovic and Weisbenner (2005)). If fund managers have superior information about stocks in particular industries, they should outperform in these industries. Kacperczyk, Sialm and Zheng (2005) show that funds with above-median industry concentration yield an average return that is 1.1% per year higher than those with below-median concentration.

The model also predicts that home investors should out-perform foreign investors on home assets. Choe, Kho and Stulz (2004) document home asset outperformance by Korean investors.

[9] On-line technical appendix D proves that concentrated portfolios achieve higher expected returns. It also uses the theory interpret measures of portfolio risk.

While one might think that this is only true for individual investors, Hau (2001) documents excess German-asset returns for *professional traders* in Germany. Similarly, Shukla and van Inwegen (1995) document that US mutual funds earn higher returns on US assets than UK funds do. Dvorak (2007) argues that Indonesian investors outperform foreigners on Indonesian assets, even when that investment is intermediated by a professional.

Cross-sectional Asset Returns Investors want to learn information others do not know because assets that many other investors learn about have high prices and low expected returns. Thus a falsifiable prediction of the model is its negative relationship between information and returns. Three studies confirm this prediction. First, Botosan (1997) and Easley, Hvidkjaer and O'Hara (2002) find that more public information lowers an asset's return. Second, Pastor and Veronesi (2003) find that firms with more abundant historical data offer lower returns. Finally, Greenstone, Oyer and Vissing-Jorgenson (2006) analyze a mandatory disclosure law that changed a group of assets from being low-information to high-information. This change should cause temporary high returns while prices are increasing, followed by lower returns going forward. They find that between proposal and passage of the law, prices of the most affected firms rose, producing abnormal excess returns of 11-22%. After passage, excess returns disappeared.

4.2 Quantitative Evaluation: Is capacity large enough?

A key unobserved variable in the model is the investor's capacity, which regulates how much he can learn. This exercise infers capacity from estimates of portfolio out-performance. The test is: Does this inferred level of capacity deliver observed home bias? This is a useful test because it tells us if home bias is rationalized by profit-maximization. Before proceeding with the main exercise, we first explore how two model assumptions affect the optimal degree of home bias: asset correlation and local information.

Two countries have 1000 identical investors each. The 5 home and 5 foreign assets are all uncorrelated. Foreigners start out α times more uncertain about home risks $(1+\alpha)\Lambda_h = \Lambda_h^{\star}$, and home investors are α times more uncertain about foreign risks $\Lambda_f = (1+\alpha)\Lambda_f^{\star}$. We consider a 10% information advantage ($\alpha = 0.1$). Risk aversion is $\rho = 2$. The supply of each asset has mean $\bar{x}$ =100 and standard deviation 10. Expected payoffs for home and foreign assets are equal. They are equally spaced between 1 and 2. The mean of the average investor's prior belief is the asset's true payoff. The standard deviation of prior beliefs is between 15-30%, such that all assets have the same prior expected payoff to standard deviation ratio. To explore various levels of capacity, we transform K into a more intuitive measure: $\tilde{K} = 1 - K^{-1/2}$ is how much an investor can reduce the standard deviation of one asset through learning. Following convention, home bias is

$$\text{home bias} = 1 - \frac{1 - share\ of\ home\ asset\ in\ home\ portfolio}{share\ of\ foreign\ assets\ in\ world\ portfolio}. \tag{11}$$

In this example, as in the data, the share of foreign assets in the world portfolio is 0.5. In a world where there is no initial information advantage and no learning capacity, home bias is zero. We use an economy with an initial information advantage, but no learning capacity as a benchmark. A 10% initial information advantage by itself generates a 5.3% home bias.

Asset Correlation Increases Home Bias With uncorrelated assets, a home investor acquires information about one home asset and over-weights that asset in his portfolio. When capacity can eliminate 22% of the standard deviation in one asset ($\tilde{K} = .22$), home bias is 10%, almost double its no-learning level. When $\tilde{K} = .70$, home bias is 45%, more than eight times larger than the home bias without learning.

Moderate correlation increases home bias because several home assets load on the one risk factor the investor learns about. When the investor has better information about more home assets, he tilts his portfolio more towards home risk. When home assets are positively correlated with each other, and foreign assets are positively correlated with each other (correlations of 10-30%), but the two sets of assets are mutually uncorrelated, home bias doubles to 19.4% for $\tilde{K} = .22$. It increases to 59.5% for $\tilde{K} = .70$. (See line with circles in figure 3.) In contrast, the no learning benchmark is unaffected (5.3%, line with diamonds). With $\tilde{K} = .82$, home bias is 72%, just shy of the 76% observed in the data. This level of capacity is still quite high. Two model features would lower required capacity: higher asset payoff correlation and advantages in local risks, which we explore next.

Local Information Increases Home Bias We use the same numerical example with correlated assets, except that instead of giving 1000 home (foreign) investors a 10% initial information advantage in all 5 home (foreign) assets, we give 200 investors each a 50% advantage in one asset;

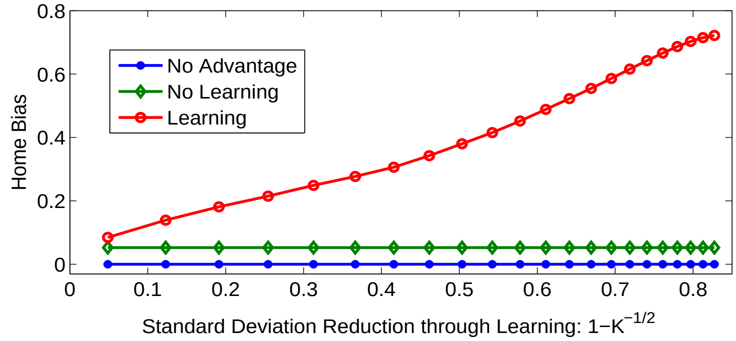

Figure 3: Home Bias Increases With Capacity. Assets within a country have correlated payoffs (cov= $.09^2$). Home bias is defined in (11). The 'no advantage' line (stars) is an economy with no initial informational advantage and no capacity to learn. The 'no learning' economy (diamonds) has a small initial information advantage (10%) and no learning capacity. The 'learning' line (circles) is our model. Learning capacity K varies between 1.1 and 30. The horizontal axis plots $\tilde{K}$, the potential percentage reduction in the standard deviation of one asset ($\tilde{K} = 1 - K^{-1/2}$).

the aggregate information advantages at home and abroad are unchanged. We measure local bias as in (11), treating localities like countries. With capacity $\tilde{K} = 0.70$, local bias is 30%. The average local investor holds 3.6 times what a diversified investor would hold, of his local asset.

Concentrating information advantages in local assets *increases home bias.* Without learning, the home bias is 8%; with low capacity ($\tilde{K} = 0.22$), it is 23%. With more capacity ($\tilde{K} = 0.70$), home bias is 76%. This is 16.5% more than in in the previous case and matches the 76% home bias in the data. The underlying capacity level K that matches the home bias in the local-advantage model is 3 times smaller than in the home country advantage model.

Inferring the level of capacity Portfolio out-performance provides clues about how much private information investors have. Ivkovic et al. (2007) use brokerage account data to show that individuals investors with concentrated portfolios earn 10% higher risk-adjusted annual returns on local, non-S&P500 stocks than investors with diversified portfolios.

To link the model to data, we equate the largest risk-factor in the home country (80% of market capitalization) with S&P500 stocks (73% of US market capitalization). For the non-S&P risk factors, we compare expected returns of local investors, who learn about the local asset, and non-local investors. For the level of capacity that matches the empirical home bias ($\tilde{K} = .70$), local investors' return on the smaller risk factors is 5% higher than what non-locals earn. The model can match Ivkovic et al. (2007)'s 10% result for ($\tilde{K} = .75$). This inference suggests that the level of capacity required to match the home bias is not implausibly large.

Ivkovic et al. (2007) focus on non-S&P500 stocks because their informational asymmetries are potentially the largest. They also report insignificant outperformance on the S&P assets. While our model cannot speak to the statistical significance of their results, it does qualitatively match the pattern of lower outperformance on larger assets. For the calibration that matches the home bias, local investors' return on the S&P risk factors is only 2% higher than what non-locals earn. Returns fall on large assets because their size makes them valuable to learn about. Low average uncertainty about the risks makes equilibrium returns and outperformance low.

4.3 Seemingly Contradictory Evidence

We discuss two facts that are inconsistent with the version of our model outlined so far. We show that both facts can be explained if we allow for asymmetric capacity (more developed financial analysis sectors in some countries than others).

Foreign Out-performance in Emerging Markets Using foreign investment data from Taiwan, Seasholes (2004) finds that foreign investors outperform the Taiwanese market, particularly in assets that are large and highly correlated with the macroeconomy. He argues that "The results point to foreigners having better information processing abilities, especially regarding macro-fundamentals." This conclusion leads us to ask two questions of our model.

Question 1: If Taiwanese investors have lower capacity than Americans, might Americans invest in Taiwanese assets and outperform the market? Recall that expected returns are determined by $\hat{\Lambda}^a$. If Americans have more capacity, they will reduce the average posterior variance for American assets by more: $\hat{\Lambda}^a_{hi} < \hat{\Lambda}^{a\star}_{fi}$, for equally-sized home and foreign risks hi and fi. Therefore, expected returns for US assets will be lower than for Taiwanese assets. A large enough difference in returns will induce some Americans to invest in Taiwan and learn about Taiwan. If Americans have capacity that exceeds Taiwanese capacity, and the capacity gap exceeds their initial disadvantage in a Taiwanese risk factor, then Americans can become the best informed of any investor about that risk factor. Being best informed, the American will out-perform the average investor in assets that load on that factor.

Question 2: Will American excess returns be concentrated in those Taiwanese assets that load heavily on the largest risk factors? Since section 4.1 shows that foreign investors learn about large assets with high market covariance, these are the Taiwanese assets American should out-perform on. Thus, an asymmetric capacity version of the model can reconcile high-capacity investors' out-performance at home, with their out-performance in emerging markets, for large high-beta assets.

The declining home bias The previous results imply that a rise in learning capacity K should increase home bias. At first glance, these results seem to suggest that home bias should increase over time. If anything, the data point to a modest decline in the U.S. home bias. However, only a symmetric increase in capacity unambiguously increases home bias. If home investors' capacity increases more, returns on home assets decline. Relatively higher foreign returns may induce some home investors to specialize in learning about and holding foreign equity. Thus, asymmetric increases in capacity could reduce the average investor's home bias.

Furthermore, capital flow liberalization and increases in equity listings in the last 30 years have increased investible foreign risk factors (Bekaert, Harvey and Lundblad 2003). The investors in our model would add these risk factors to the 'no advantage' part of their portfolio ($\bar{x}$). This effect would also increase foreign equity investment and reduce home bias.

4.4 A New Direction for Estimating Information

The fact that investors' information is inherently unobservable is an obstacle to assessing asymmetric information theories. One solution is to use proxies for investors' information, like the precision of earnings forecasts. But for many classes of investors, such proxies are not available. Our theory offers another solution. It delivers information sets as equilibrium outcomes. Observable features of assets predict information patterns, which in turn, predict observable portfolios, analyst behavior and pricing errors. This makes for testable hypotheses. A contribution of this paper is that it brings information-based theories to the data.

The novel part of this theory is the link it establishes between observable asset characteristics and the average investor's information, through the learning index. The following algorithm could be used to estimate learning indices: (i) Compute the eigen-decomposition (principal components) of asset payoffs. Payoffs are the dividend paid between t and $t+1$ plus the price at $t+1$: $f_t = d_t + p_{t+1}$. Post-multiply asset prices and payoffs by the eigenvector matrix Γ, to form risk factor prices and payoffs. Risk factor returns are $(f_t - rp_t)\Gamma$. (ii) Construct unconditional (prior) risk factor Sharpe ratios: Divide each risk factor's average return by its standard deviation. (iii) Estimate the coefficient Λ_B from a regression of risk factor prices ($\Gamma' p$) on a constant and risk factor payoffs ($\Gamma' f$)

– the risk factor counterpart to the price equation (12). Attributing the residual to asset supply shocks, the residual variance is $\Lambda_C^2\sigma_x^2$. One minus that regression's R^2 is Λ_{pi}/Λ_i for an investor whose prior belief is based on past realizations of returns.[10] (iv) Add the squared Sharpe ratio to $(1-R^2)$ to obtain the learning index for each risk factor. (v) Pre-multiply the vector of risk factor indices by the eigenvector matrix Γ. The resulting vector contains learning indices for each asset. Alternatively, this procedure could be applied to countries or regions by using market indices for prices and returns.

Learning indices could be used to test many aspects of the theory. (1) They should predict information-related variables such as analyst coverage. (2) Countries, regions or firms with higher learning indices should have lower returns, relative to what a standard model like the CAPM predicts. This is because the average investor is less uncertain about an asset he learns more about (higher learning index), and because lower uncertainty implies a lower return. When the average investor learns more about an asset with a higher index, he reduces its risk and therefore its return. (3) Finally, a country or region's learning index should be related to the home bias of its residents' portfolios. This relationship is non-monotonic. If the learning index is near zero, no one, not even locals learn about home risk. When all investors learn about foreign risk, there is only a small home bias that comes from initial information differences. As the home learning index grows, more home investors specialize in home risks. Information asymmetry and home bias rise. In the limit, as the home learning index grows very large, all investors study home risks. Again, the small home bias comes only from the small differences in initial information. Because home bias depends on comparative information advantage, it is strongest for an intermediate level of the learning index.

5 Conclusions

Every expectations operator, every variance and every covariance is conditioned on an information set. Therefore, every asset pricing and portfolio choice model makes assumptions about what information agents use. Most theories employ stylized informational assumptions: complete information about all past events and no additional information about future events. Their predictions rest on these assumptions. While preferences have been re-examined and new risk factors have been

[10]To derive the link between the regression R^2 and the learning index, manipulate (14) to get $C = -\rho\Sigma_\eta^a$. Square this equation and use (16) to get $C^2\sigma_x^2 = \Sigma_p$. Since $C^2\sigma_x^2$ is the unexplained sum of squares in the price regression and Σ is the total variance in prices, the regression $(1-R^2)$ is $\Sigma^{-1}\Sigma_p$, for assets and Λ_{pi}/Λ_i for risk factor i.

identified, relatively little work has explored information sets. We can't observe information, but we can ask what investors would observe if they had a choice. Predicting information sets on the basis of observed features of assets circumvents the problem of unobserved information.

This paper questions the common assumption that residents have more information about their region's assets than do non-residents. If investors are restricted in the amount of information they can learn about risky asset payoffs, which assets would they choose to learn about? Investors who do not account for the effect of learning on portfolio choice, choose to undo their initial advantages. But, investors with rational expectations reinforce informational asymmetries. Investors learn more about risks they have an advantage in because they want their information to be very different from what others know. Thus our main message is that information asymmetry assumptions are defensible, but not for the reason originally thought. We do not need cross-border information frictions. With sufficient capacity to learn, small initial information advantages can lead to a home bias of the magnitude observed in the data.

An important assumption in our model is that every investor must process his own information. But paying one portfolio manager to learn for many investors is efficient. How might such a setting regenerate a home bias? Because monitoring information collection is difficult, portfolio managers have an incentive to lie about how much research they do. Investors may want to occasionally audit portfolio managers. Thus having a manager from the same region, with similar initial information, is advantageous because checking the manager's work requires less capacity. Portfolio managers with the same initial information advantage as their clients form the same optimal portfolio as would a client who processed information himself. This optimal portfolio is home biased. Future work could use the framework in this model to build an equilibrium model of delegated portfolio management and investigate the effect of portfolio managers on asset prices as in Cuoco and Kaniel (2006).

A broader message of our paper is that investors choose to have different information sets. Models that assume symmetric information, commonly used in finance, are subject to a criticism: Investors have an incentive to deviate by learning information that others do not know.

References

Admati, Anat, "A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets," *Econometrica*, 1985, *53(3)*, 629–57.

—— **and Paul Pfleiderer**, "Direct and Indirect Sale of Information," *Econometrica*, 1990, *58(4)*, 901–28.

Ahearne, Alan, William Griever, and Francis Warnock, "Information Costs and Home Bias: An Analysis of US Holding of Foreign Equities," *Journal of International Economics*, 2004, *62*, 313–36.

Bacchetta, Philippe and Eric van Wincoop, "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?," *American Economic Review*, 2006, *96(3)*, 552–576.

Bae, Kee-Hong, Rene Stulz, and Hongping Tan, "Do Local Analysts Know More? A Cross-Country Study of the Performance of Local Analysts and Foreign Analysts," 2005. NBER Working Paper 11697.

Banerjee, Snehal, "Learning from Prices and the Dispersion in Beliefs," 2007. Stanford GSB working paper.

Bekaert, Geert, Campbell Harvey, and Christian Lundblad, "Equity Market Liberalization in Emerging Markets," *Journal of Financial Research*, 2003, *26*, 275–299.

Botosan, Caval, "Disclosure Level and the Cost of Equity Capital," *The Accounting Review*, 1997, *72*, 323349.

Boyle, Phelim, Raman Uppal, and Tan Wang, "Ambiguity Aversion and the Puzzle of Own-Company Stock in Pension Plans," 2003. Working Paper.

Brennan, Michael and Henry Cao, "International Portfolio Investment Flows," *The Journal of Finance*, 1997, *52*, 1851–80.

Calvo, Guillermo and Enrique Mendoza, "Rational Contagion and the Globalization of Securities Markets," *Journal of International Economics*, 2000, *51(1)*, 79–113.

Choe, Hyuk, Bon-Sang Kho, and Rene Stulz, "Do Domestic Investors Have an Edge? The Trading Experience of Foreign Investors in Korea," 2004. Working Paper.

Cohen, Lauren, "Loyalty Based Portfolio Choice," 2004. Working Paper.

——**, Andrea Frazzini, and Christopher Malloy**, "The Small WOrld of Investing: Board Connections and Mutual Fund Returns," 2007. NBER Working Paper #13121.

Cole, Harold, George Mailath, and Andrew Postlewaite, "Investment and Concern with Relative Position," *Review of Economic Design*, 2001, *6(2)*, 241–261.

Coval, Joshua, "International Capital Flows when Investors Have Local Information," 2000. Working Paper.

—— **and Tobias Moskowitz**, "The Geography of Investment: Informed Trading and Asset Prices," *Journal of Political Economy*, 2001, *109(4)*, 811–41.

Cover, Thomas and Joy Thomas, *Elements of Information Theory*, first ed., John Wiley and Sons, New York, New York, 1991.

Cuoco, Domenico and Ron Kaniel, "Equilibrium Prices in the Presence of Delegated Portfolio Management," 2006. Working Paper.

DeMarzo, Peter, Ron Kaniel, and Ilan Kremer, "Diversification as a Public Good: Community Effects in Portfolio Choice," *The Journal of Finance*, 2004, *59(4)*, 1677–1715.

Dow, James, Itay Goldstein, and Alexander Guembel, "Incentives for Information Production in Markets where Prices Affect Real Investment," 2007. Working Paper.

Dvorak, Tomas, "Do Domestic Investors Have an Information Advantage? Evidence from Indonesia," *The Journal of Finance*, 2007, *60(2)*, 817–839.

Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, "Is Information Risk a Determinant of Asset Returns?," *The Journal of Finance*, 2002, *57(5)*, 2185–2221.

Evans, Martin and Richard Lyons, "A New Micro Model of Exchange Rate Dynamics," 2004. Working Paper.

French, Kenneth and James Poterba, "International Diversification and International Equity Markets," *American Economic Review*, 1991, *81(2)*, 222–26.

Glodjo, Arman and Campbell Harvey, "Forecasting Foreign Exchange Market Movements via Entropy Coding," 1995. Working Paper.

Graham, John, Campbell Harvey, and Hai Huang, "Investor Competence, Trading Frequency, and Home Bias," 2006. Working Paper.

Greenstone, Michael, Paul Oyer, and Aneette Vissing-Jorgenson, "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments," *The Quarterly Journal of Economics*, 2006, *121*, 2399–460.

Grossman, Sanford and Joseph Stiglitz, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 1980, *70(3)*, 393–408.

Guiso, Luigi and Tullio Jappelli, "Information Acquisition and Portfolio Performance," 2006. CEPR Working Paper #5901.

Hansen, Lars and Thomas Sargent, "Robust Control and Model Uncertainty," *American Economic Review*, 2001, *91(2)*, 60–66.

Hau, Harald, "Location Matters: An Examination of Trading Profits," *Journal of Finance*, 2001, *56(5)*, 1959–83.

Huberman, Gur, "Familiarity Breeds Investment," *Review of Financial Studies*, 2001, *14(3)*, 659–80.

Ivkovic, Zoran and Scott Weisbenner, "Local Does as Local Is: Information Content and the Geography of Individual Investors' Common Stock Investments," *Journal of Finance*, 2005, *60*, 267–306.

—— **, Clemens Sialm, and Scott Weisbenner**, "Portfolio Concentration and the Performance of Individual Investors," *Journal of Financial and Quantitative Analysis*, 2007.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, "On the Industry Concentration of Actively Managed Equity Mutual Funds," *Journal of Finance*, 2005, *60(4)*, 1983–2011.

Kang, Jun-Koo and Rene Stulz, "Why Is There a Home Bias? An Analysis of Foreign Equity Ownership in Japan," *Journal of Financial Economics*, 1997, *46(1)*, 3–28.

Lewis, Karen, "Trying to Explain Home Bias in Equities and Consumption," *Journal of Economic Literature*, 1999, *37*, 571–608.

Malloy, Christopher, "The Geography of Equity Analysis," *Journal of Finance*, 2005, *60*, 719–755.

Massa, Massimo and Andrei Simonov, "Hedging, Familiarity and Portfolio Choice," *The Review of Financial Studies*, 2006, *19(2)*, 633–685.

Morse, Adair and Sophie Shive, "Patriotism in Your Portfolio," 2003. Working Paper.

Moscarini, Giuseppe, "Limited Information Capacity As A Source of Inertia," *Journal of Economic Dynamics and Control*, 2004, *28(10)*, 2003–35.

Ozdenoren, Emre and Kathy Yuan, "Feedback Effects and Asset Prices," 2007. Working Paper.

Pastor, Lubos, "Portfolio Selection and Asset Pricing Models," *Journal of Finance*, 2000, *55(1)*, 179–223.

—— **and Pietro Veronesi**, "Stock Valuation and Learning About Profitability," *Journal of Finance*, 2003, *58(5)*, 1749–89.

Peress, Joel, "The Tradeoff between Risk Sharing and Information Production in Financial Markets," 2006. INSEAD Working Paper.

Portes, Richard, Helene Rey, and Yonghyup Oh, "Information and Capital Flows: The Determinants of Transactions in Financial Assets," *European Economic Review*, 2001, *45*, 783–96.

Seasholes, Mark, "Re-examining Information Asymmetries in Emerging Stock Markets," 2004. Working Paper.

Shannon, Claude E., "Mathematical Theory of Communication," *Bell System Technology Journal*, 1948, *27*, 379–423 and 623–656.

Shukla, Ravi and Greg van Inwegen, "Do Locals Perform Better than Foreigners?: An Analysis of UK and US Mutual Fund Managers," *Journal of Economics and Business*, 1995, *47(3)*, 241254.

Sims, Christopher, "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50(3)*, 665–90.

Tesar, Linda and Ingrid Werner, "The Internationalization of Securities Markets Since the 1987 Crash," 1998. Brookings-Wharton Papers on Financial Services, The Brookings Institution.

Uppal, Raman and Tan Wang, "Model Misspecification and Underdiversification," *Journal of Finance*, 2003, *58(6)*, 2465–2486.

Van Nieuwerburgh, Stijn and Laura Veldkamp, "Information Acquisition and Under-Diversification," 2006. Working Paper.

Yuan, Kathy, "Asymmetric Price Movements and Borrowing Constraints: A REE Model of Crisis, Contagion, and Confusion," *Journal of Finance*, 2005, *60(1)*, 379–411.

A Proofs

A.1 Proof of Proposition 1

Step 1: Derive the optimal learning strategy Claim: Optimal learning about principal components Γ produces a posterior belief $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma$ with eigenvalues $\hat{\Lambda}_i = \min(\Lambda_i, \frac{1}{\tilde{q}_i^2}M)$, where M is a constant, common to all assets.
Proof: The optimization problem is

$$\max_{\hat{\Lambda}} \sum_i \tilde{q}_i^2 \hat{\Lambda}_i$$

s.t. $\hat{\Lambda}_i \leq \Lambda_i$ and $\prod_i \hat{\Lambda}_i \geq \prod_i \Lambda_i \frac{1}{K}$, where $\tilde{q}_i = \Gamma_i' q$. The first-order condition for this problem is

$$\tilde{q}_i^2 - \upsilon \frac{1}{\hat{\Lambda}_i} \prod_i \hat{\Lambda}_i + \phi_i = 0$$

where υ is the Lagrange multiplier on the capacity constraint and ϕ_i is the Lagrange multiplier on the no-negative-learning constraint for asset i. Define $M = \upsilon \frac{1}{K} \prod_i \Lambda_i$. The result that $\hat{\Lambda}_i = \min\left\{\Lambda_i, \frac{M}{\tilde{q}_i^2}\right\}$ follows from the first order condition and the no-negative learning constraint, which states that $\phi_i = 0$ when $\hat{\Lambda}_i > \Lambda_i$.

Step 2: Show that learning eliminates initial advantages If $(\Lambda_i - \epsilon)\tilde{q}_i^2 > M$ for $i = argmin_j (\Lambda_j - \epsilon)\tilde{q}_j^2$, then the optimal learning strategy tells us that posterior beliefs $\hat{\Lambda}_i$ are unaffected by an ϵ reduction in the prior belief. There exists a capacity $K^\star$ such that $\min_i \left((\Lambda_i - \epsilon)\tilde{q}_i^2\right) = M$. All that is left is to characterize $K^\star$. Since the capacity used learning about a factor j is $\Lambda_j/\hat{\Lambda}_j) = \Lambda_j/(\frac{1}{\tilde{q}_j^2}M) = (\Lambda_j\tilde{q}_j^2)/\left(\min_i \left((\Lambda_i - \epsilon)\tilde{q}_i^2\right)\right)$, the total capacity required is

$$K^\star = -\left(\min_i \left((\Lambda_i - \epsilon)\tilde{q}_i^2\right)\right)^N \left(\prod_{j=1}^N \Lambda_j \tilde{q}_j^2\right).$$

A.2 The required price of foreign information without increasing returns

Consider a setting with one home and one foreign asset, with prior variances σ_h^2 and σ_f^2, posterior variances $\hat{\sigma}_h^2$ and $\hat{\sigma}_f^2$, and zero covariance. Replace (3) with a capacity constraint that requires ψ times more capacity to process foreign than home information: $\sigma_h/\hat{\sigma}_h \cdot (\sigma_f/\hat{\sigma}_f)^{\psi} \leq K$. The optimal learning choice is described by the $\hat{\sigma}_h^2$ and $\hat{\sigma}_f^2$ first order conditions. Capacity permitting, an investor sets the ratio of posterior variances to $\hat{\sigma}_f^2/\hat{\sigma}_h^2 = \psi q_h^2/q_f^2$. The investor's optimal portfolio is: $q^\star = 1/\rho\widehat{\Sigma}^{-1}(\hat{\mu} - pr)$. This implies that an investor, who initially expects to hold a balanced portfolio ($q_h = q_f$) but ends up holding 7.3 times more home assets (as in the data), must have $\psi = 7.3$. If home and foreign expected returns ($\hat{\mu} - pr$) are equal and covariance is zero, then $q_h^\star/q_f^\star = \hat{\sigma}_f^2/\hat{\sigma}_h^2$.

The zero covariance assumption biases ψ downward by overestimating home bias in two ways. (1) It makes gains to diversification large. (2) If home signals are informative about correlated foreign assets, home bias would fall and the required cost differential would have to be even higher.

Adding an initial home advantage does not alter this required processing cost, unless the advantage alone can account for the home bias. Of course, home bias could arise if an investor anticipated holding lots of home assets: $q_h > q_f$. But then home bias comes not from processing costs, but from portfolio expectations. This is the mechanism explored in section 3.

A.3 Equilibrium Asset Prices

From Admati (1985), we know that equilibrium price takes the form

$$rp = A + Bf + Cx \quad \text{where} \tag{12}$$

$$A = -\rho\left(\frac{1}{\rho^2\sigma_x^2}(\Sigma_\eta^a\Sigma_\eta^{a'})^{-1} + (\Sigma_\eta^a)^{-1}\right)^{-1}\bar{x}, \tag{13}$$

$$C = -\left(\frac{1}{\rho^2\sigma_x^2}(\Sigma_\eta^a\Sigma_\eta^{a'})^{-1} + (\Sigma_\eta^a)^{-1}\right)^{-1}\left(\rho I + \frac{1}{\rho\sigma_x^2}(\Sigma_\eta^a)^{-1'}\right). \tag{14}$$

The matrix B is the identity matrix, because all investors have independently distributed priors. We treat priors as though they were private signals. This assumption deviates from Admati (1985) and Van Nieuwerburgh and Veldkamp (2004), which assumes that investors have identical priors.

Let $\Sigma_{\eta j}$ be the variance-covariance matrix of the private signals that investor j chooses to observe. For future use, we define the following three precision matrices. They are derived from the above pricing function and the definitions for A, B, and C. $(\Sigma_\eta^a)^{-1}$ is the average precision of investors' information advantage, plus the average precision of the information they choose to learn. $(\Sigma_p)^{-1}$ is the precision of prices as a signal about true payoffs. $(\widehat{\Sigma}^a)^{-1}$ is the average of all investors' posterior belief precisions, taking into account priors, signals and prices.

$$(\Sigma_\eta^a)^{-1} = \Gamma(\Lambda_\eta^a)^{-1}\Gamma' = \frac{1}{2}\Sigma^{-1} + \frac{1}{2}(\Sigma^\star)^{-1} + \int_j (\Sigma_\eta^j)^{-1} dj, \tag{15}$$

$$(\Sigma_p)^{-1} = \Gamma\Lambda_p^{-1}\Gamma' = \frac{1}{\rho^2\sigma_x^2}(\Sigma_\eta^a\Sigma_\eta^{a'})^{-1}, \tag{16}$$

$$(\widehat{\Sigma}^a)^{-1} = \Gamma\hat{\Lambda}_a^{-1}\Gamma' = \frac{1}{\rho^2\sigma_x^2}(\Sigma_\eta^a\Sigma_\eta^{a'})^{-1} + (\Sigma_\eta^a)^{-1} \tag{17}$$

We have assumed that investors choose to obtain signals about the eigenvectors Γ of the prior covariance matrix Σ. It is easy to show that when Σ_η has eigenvectors Γ, the three precision matrices above also have the same eigenvectors.

We note and later use that $CC'\sigma_x^2 = \rho^2\sigma_x^2\Sigma_\eta^a\Sigma_\eta^{a\prime} = \Sigma_p$, because $C = -\rho\Sigma_\eta^a$. We also use that expected risk factor returns are

$$\Gamma_i'E[f - pr] = -\Gamma_i'A = \rho\Gamma_i'\hat{\Sigma}^a\bar{x} = \rho\Gamma_i'\Gamma\hat{\Lambda}^a\Gamma'\bar{x} = \rho(\Gamma_i'\bar{x})\hat{\Lambda}_i^a, \tag{18}$$

where the first equality follows from the definition of A and the definition of $\hat{\Sigma}^a$, the second equality follows from $\hat{\Sigma}^a = \Gamma\hat{\Lambda}^a\Gamma'$, and the last equality follows from $\Gamma'\Gamma = I$.

A.4 Proof of Proposition 2

Expected excess returns $\hat{\mu} - pr$ are normally distributed with mean $-A$ and variance $V_{ER} = \Sigma_p - \widehat{\Sigma}$. The first part of the objective is $Tr\left(\widehat{\Sigma}^{-1} V_{ER}\right)$, which we rewrite as $Tr\left(\widehat{\Sigma}^{-1}\Sigma\Sigma^{-1}(V_{ER} + \widehat{\Sigma} - \widehat{\Sigma})\right)$. This is $Tr\left(\widehat{\Sigma}^{-1}\Sigma\Sigma^{-1}(V_{ER} + \widehat{\Sigma}) - I\right)$ or $Tr\left(\widehat{\Sigma}^{-1}\Sigma\Sigma^{-1}(V_{ER} + \widehat{\Sigma})\right) - N$. The trace is the sum of the eigenvalues. Let y_i, be the ratio of the precision of the posterior to the precision of the prior for risk i, i.e. it is the i^{th} eigenvalue of $\widehat{\Sigma}^{-1}\Sigma$: $y_i \equiv \hat{\Lambda}_i^{-1}\Lambda_i$. Let X_i be the i^{th} eigenvalue of $\Sigma^{-1}(V_{ER} + \widehat{\Sigma})$. Then the i^{th} eigenvalue of the matrix inside the trace is $y_i X_i$, and $Tr\left(\widehat{\Sigma}^{-1}\Sigma\Sigma^{-1}(V_{ER} + \widehat{\Sigma})\right) = \sum_{i=1}^{N} X_i y_i$. This is because Σ, $\widehat{\Sigma}$, and C all share the same eigenvectors Γ. The matrix $\Sigma^{-1}(V_{ER} + \widehat{\Sigma}) = \Sigma^{-1}\Sigma_p$ has eigenvalues $X_i = \Lambda_{pi}\Lambda_i^{-1}$. The second part of the object function is $\sum_{i=1}^{N} \theta_i^2 y_i$, where $\theta_i^2 = (\Gamma_i' A)^2 \Lambda_i^{-1}$ is the prior squared Sharpe ratio of risk factor i. The objective is to maximize $\sum_{i=1}^{N}(X_i + \theta_i^2) y_i$, where $X_i + \theta_i^2$ is the *learning index* of risk factor i. The maximization over $\{y_i\}$ is subject to $\prod_{i=1}^{N} y_i \leq K$ and $y_i \geq 1 + \frac{\Lambda_{pi}^{-1}}{\Lambda_i^{-1}}$. This problem maximizes a sum subject to a product constraint. A simple variational argument shows that the maximum is attained by maximizing the y_i with the highest learning index $X_i + \theta_i^2$. The investor devotes all his 'spare capacity' to learning about this risk factor i. To be more precise, he sets $y_j = 1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}$, for all risk factors j that he does not learn about, and he uses all remaining capacity to obtain a private signal on risk factor i: $y_i = \tau\left(1 + \frac{\Lambda_{pi}^{-1}}{\Lambda_i^{-1}}\right)$, where $\tau = K\left(\prod_{j=1}^{N}\left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right)\right)^{-1}$. We endow the investor with enough capacity such that he has spare capacity to acquire private signals after devoting capacity to learning from prices: $\prod_{j=1}^{N}\left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right) < K$ and therefore $\tau > 1$. For future reference define the 'spare capacity' of an investor who learns about risk factor i as

$$\tilde{K}_i = \prod_{j \neq i}\left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right)^{-1}. \tag{19}$$

A.5 Proof of Proposition 3

The learning index for home risk factor i is always greater for a home investor:

$$\frac{\Lambda_{pi}}{\Lambda_i} + \frac{(\hat{\Lambda}_i^a)^2}{\Lambda_i}(\Gamma_i'\bar{x})^2 > \frac{\Lambda_{pi}}{\Lambda_i^\star} + \frac{(\hat{\Lambda}_i^a)^2}{\Lambda_i^\star}(\Gamma_i'\bar{x})^2. \tag{20}$$

because $\Lambda_i < \Lambda_i^\star$. Likewise, the learning index of a foreign risk factor j is always greater for a foreign investor:

$$\frac{\Lambda_{pj}}{\Lambda_j^\star} + \frac{(\hat{\Lambda}_j^a)^2}{\Lambda_j^\star}(\Gamma_j'\bar{x})^2 > \frac{\Lambda_{pj}}{\Lambda_j} + \frac{(\hat{\Lambda}_j^a)^2}{\Lambda_j}(\Gamma_j'\bar{x})^2. \tag{21}$$

because $\Lambda_j > \Lambda_j^\star$.

Therefore, if one foreign investor learns about a home risk factor i, then all home investors must also be learning about i, or some other risk factor with an equally high learning index. This other risk factor must be a home risk factor, otherwise the foreign investor would strictly prefer to learn about it. Let $\tilde{K}$ be the spare capacity for a particular home risk factor (19). Since every home investor learns about that home risk factor, then $|\hat{\Lambda}_h| = \frac{1}{\tilde{K}}|\Lambda_h^{-1} + \Lambda_p^{-1}|^{-1}$ and $|\hat{\Lambda}_f| = |\Lambda_f^{-1} + \Lambda_p^{-1}|^{-1}$. Since foreign investors might learn about home risk, but might not: $|\hat{\Lambda}_h^\star| \geq \frac{1}{\tilde{K}}|\Lambda_h^\star + \Lambda_p^{-1}|^{-1}$ and since he might or might not learn about his own foreign risk: $|\hat{\Lambda}_f^\star| \leq |\Lambda_f^\star + \Lambda_p^{-1}|^{-1}$. Since price precisions (Λ_p^{-1}) are constant and positive, taking ratios of $|\hat{\Lambda}_h^\star|$ to $|\hat{\Lambda}_h|$ and of $|\hat{\Lambda}_f|$ to $|\hat{\Lambda}_f^\star|$ yields the result. The same argument can be made, in the case where one or more home investors learn about foreign risks.

A.6 Proof of Proposition 4

Symmetric risk factors From the previous proposition, we know that an investor with $K > 0$ will learn about a risk factor that they have an advantage in, one of their home risk factors. Let i denote that risk factor. Then $\hat{\Lambda}_i^{-1} = K\Lambda_i^{-1}$. When investors can learn ($K > 0$), let ξ_i denote the fraction of home investors that learn about home risk factor i. Then $(\hat{\Lambda}_i^a)^{-1} = \frac{1}{2}\xi_i K(\Lambda_i)^{-1} + \frac{1}{2}(1-\xi_i)(\Lambda_i)^{-1} + \frac{1}{2}(\Lambda_i^\star)^{-1}$. The product $\hat{\Lambda}_i^{-1}\hat{\Lambda}_i^a$ is increasing in K because the first term is increasing proportionally and the second term is decreasing less than proportionally in K. Using equation (10), describing the portfolio with $K > 0$ and the no learning portfolio ($K = 0$), it follows that the difference between the i^{th} component, $\Gamma_i'(\hat{\Lambda}^{-1}\hat{\Lambda}^a - \Lambda^{-1}\Lambda^a)(\Gamma\bar{x})$ is strictly positive.

General case From the previous proposition, we know there are three situations to consider: all investors learn about their own home assets, some home investors learn about foreign risk factors, or some foreigners learn about home risk factors. This first case we considered in the previous paragraph. We prove the third case here; the second one follows from the same logic.

When some foreign investors learn about home risks, all home investors must learn about home risks as well. Every investor who learns about home risks is indifferent between learning about any home risk learned about in equilibrium. While the extent of home bias won't hinge on which risk factor, within a country, any investor learns about, it simplifies our analysis to assume that each investor who learns about home risks adopts a symmetric mixed strategy over which risks to specialize in. Let ξ_i ($\xi_i^\star$) be the fraction of home (foreign) investors who learn about home risk i. Because all home investors learn about home risks, it must be that: $\xi_i \geq \xi_i^\star$.

Define $\Gamma_i' q^{ha} = (\hat{\Lambda}_i^{-1})^{ha}\hat{\Lambda}^a\Gamma_i'\bar{x}$ to be the portfolio holdings of risk factor i of the average home investor (ha). This follows from pre-multiplying both sides of equation (10) by Γ_i'. Here, $(\hat{\Lambda}_i^{-1})^{ha} = \xi_i K(\Lambda_i)^{-1} + (1-\xi_i)(\Lambda_i)^{-1}$ is the average posterior precision of home investors about risk factor i. The worldwide average precision is $(\hat{\Lambda}_i^a)^{-1} = \frac{1}{2}\xi_i K(\Lambda_i)^{-1} + \frac{1}{2}(1-\xi_i)(\Lambda_i)^{-1} + \frac{1}{2}\xi_i^\star K(\Lambda_i^\star)^{-1} + \frac{1}{2}(1-\xi_i^\star)(\Lambda_i^\star)^{-1}$.

Consider the extreme case where all foreign investors learn about home risk factors ($\xi_i = \xi_i^\star$). Then $(\hat{\Lambda}_i^{-1})^{ha}\hat{\Lambda}_i^a$ can be shown to collapse to $\frac{2\hat{\Lambda}_i^{-1}}{\hat{\Lambda}_i^{-1}+(\hat{\Lambda}_i^\star)^{-1}}$. This expression does not depend on K. This implies that the learning portfolio ($K > 0$) and the no-learning portfolio ($K = 0$) are identical: $E[q_i] = E[q_i^{no\ learn}]$.

In all other cases, $\xi_i > \xi_i^\star$. Taking a partial derivative of $(\hat{\Lambda}_i^{-1})^{ha}\hat{\Lambda}^a$ reveals that it is increasing in K. As a result, the difference between the learning and the no-learning portfolio on risk factor i is strictly positive: $E[q_i] > E[q_i^{no\ learn}]$.