Information Immobility and the Home Bias Puzzle

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

July 31, 2005

Abstract

Many explanations for home or local bias rely on information asymmetry: investors know more about their home assets. A criticism of these theories is that asymmetry should disappear when information is tradable. This criticism is flawed. If investors have asymmetric prior beliefs, but choose how to allocate limited learning capacity before investing, they will not necessarily learn foreign information. Investors want to exploit increasing returns to specialization: The bigger the home information advantage, the more desirable are home assets; but the more home assets investors expect to own, the higher the value of additional home information. Even with a tiny home information advantage, and even when foreign information is no harder to learn, many investors will specialize in home assets, remain uninformed about foreign assets, and amplify their initial information asymmetry. The more investors can learn, the more home biased their portfolios become. The model's predictions are consistent with observed patterns of foreign investment, returns, and portfolio flows.

*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 Dave Backus, Chris Sims, Eric Van Wincoop, Mark Wright, and seminar participants at the Prague workshop in macroeconomic theory, the Budapest SED, the CEPR Asset Pricing meetings in Gerzensee, Ohio State, Illinois, Iowa, Princeton, Virginia, GWU and NYU, and to our anonymous referees and the Editor for helpful comments. JEL classification: F30, G11, D82. Keywords: Home bias, asymmetric information, information theory.

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. An American investor, for example, could have a highly diversified portfolio simply by purchasing foreign stocks or ADRs on US exchanges. Another hypothesis is that investors have superior access to information about local firms or economic conditions (Brennan and Cao 1997, Hatchondo 2004). But this seems to replace the assumption of capital immobility with the equally implausible assumption of information immobility. For example, if an American wished, she could presumably pay someone to divulge information about foreign firms. Such trade in information could potentially undermine the home bias.

We nevertheless propose information as an explanation for home bias. The question to be addressed, then, is why information does not flow freely across borders. Using tools from information theory (Sims 1998, 2003), we model an investor who faces a choice about what to learn, before forming his portfolio. This investor will naturally build on his existing advantage in local information because there are increasing returns to specializing in learning about one asset. A small information advantage makes a local asset less risky to a local investor. Therefore, he expects to hold slightly more local assets than a foreign investor would. But, 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. The initial small information advantage is magnified. The result is that information market segmentation persists not because investors can't learn what locals know, nor because it is too expensive, but because they don't choose to; capitalizing on what they already know is a more profitable strategy. Information immobility is plausible because information is a good with increasing returns.

In section 2, we argue that an initial information advantage alone is not enough to generate the home bias. To make this point, we examine a model where the increasing returns to learning mechanism is shut down by forcing investors to take their portfolios as given, when they choose what to learn. These investors minimize investment risk by learning about risk factors that they are most uncertain about. With sufficient capacity learning undoes all initial information advantage, and therefore all home bias. To generate a large home bias, the cost of processing foreign information would have to be larger than what is implied by the data.

Section 3 describes a general equilibrium, rational expectations model where investors choose what home or foreign information to learn, and then choose what assets to hold. The interaction of the information decision and the portfolio decision causes investors to learn information that magnifies their initial advantage. Consider two possible learning and investment strategies. One strategy would be to learn a small amount about every asset. Small changes in beliefs about every asset's payoff would cause small deviations from a diversified portfolio. Another strategy would be to learn as much as possible about a small number of assets, and then take a large position in those assets. A portfolio biased toward well-researched assets poses less risk, because a large fraction of the portfolio has been made substantially less risky, through learning. Efficient learning dictates that investors should specialize. They should learn about assets they already know well, amplify their initial information differences, and increase their home bias.

It is not the information constraint that drives investors to specialize. The model in section 2 uses the same constraint, yet investors who take portfolios as given want to equalize uncertainty across risks. Rather, it is the feedback of the learning choice and the portfolio choice on each other that generates the increasing returns. The feedback arises from the unique properties of information as a good: the more shares a piece of information can be applied to, the more benefit it provides. This idea dates back to Wilson (1975), who found that information value is increasing in a firm's scale of operation. Because of this property, information has increasing returns in many settings (Radner and Stiglitz 1984).

Calvo and Mendoza (2000) argue that more scope for diversification decreases the *incentive to learn*. In contrast, our paper shows that when investors can choose what to learn about, the *incentive to diversify* declines. Optimal portfolios contain a diversified component plus assets that the investor learns about. Equilibrium asset returns induce investors to take a long position in the assets they learn about, on average. Asset returns reflect the risk that the average investor bears. An investor who specializes in home assets becomes more informed than the average investor and earns *excess risk-adjusted home returns*. To capture, the excess return, investors take positive positions in their home assets. With higher capacity, the investor holds a larger learning component, diversification falls, and the home bias increases.

A numerical example (section 3.5) shows that learning can magnify the home bias considerably. When all home investors get a small initial advantage in all home assets (10% lower variance), the home bias is between 5 and 46%, depending on the magnitude of investors' learning capacity. When each home investor gets a local advantage, that is concentrated in one local asset, the home bias rises as high as the 76% home bias in U.S. portfolio data.

A variety of evidence supports the model's predictions. First, locally-biased portfolios earn higher abnormal returns on local stocks than more diversified ones (Coval and Moskowitz, 2001; Ivkovic, Sialm and Weisbenner, 2004). Section 4.1 shows that in a model where investors have slightly more prior information about their region, they hold more local assets and earn abnormal returns on those assets. Second, foreigners invest primarily in large stocks that are highly correlated with the market (Kang and Stultz, 1997) and often outperform locals in these assets (Seasholes, 2004). Section 4.2 shows that a foreigner with more learning capacity than locals may learn about a local risk factor. The optimal risk to learn will be one that the largest assets load on. With more information than the average investor, he will outperform the market for the assets which load on the factor: large assets that covary highly with other large assets. Third, nearby markets with highly correlated returns, generate abundant information flows (Portes and Rey, 2003), large gross equity flows (Portes, Rey and Oh, 2001) and low turnover rates (Tesar and Werner, 1994). Section 4.3 argues that having a home advantage in risks that nearby countries share, operates like having a neighboring country advantage as well. This advantage makes learning about the neighbor more profitable, and makes trading with a neighbor more like trading with a compatriot.

Magnifying information advantages generates effects that resemble a familiarity bias (Huberman 2001, Hong, Kubik and Stein 2004) or a loyalty effect (Cohen 2004). Massa and Simonov (2004) argue that familiarity effects are information driven. They find that familiarity affects less-informed investors more, diminishes when the profession or location of the investor changes, and generates higher returns.

The information choices we investigate are similar to those in models of rational inattention. However, that work has focused on time-series phenomena: delayed response to shocks (Sims 2003), inertia (Moscarini 2004), time to digest (Peng and Xiong 2005), and consumption smoothing (Luo 2005). Instead, we relax the representative agent assumption and focus on the cross-section of individuals' learning choices. Using a framework similar to Van Nieuwerburgh and Veldkamp (2005), we introduce a two-country structure. The initial information differences allow us to explore what home investors learns and what assets they hold.

Information advantages have been used to explain exchange rate fluctuations (Evans and Lyons, 2004, Bacchetta and van Wincoop, 2004), 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). All of these explanations are bolstered by our finding that information advantages are not only sustainable when information is mobile, but that asymmetry is often amplified when investors can choose what to learn.

1 A Model of Learning and Investing

We begin by setting up a general framework in which to think about learning and investment choices. In section 2 we examine the choice of what to learn when an investor takes his portfolio as given and only wants to reduce the risk of that portfolio. Section 3 describes how learning and investment decisions are made jointly in a noisy rational expectations, general equilibrium model.

This is a static model which we break up into 3-periods. In period 1, a continuum of investors choose the distribution from which to draw signals about the payoff of the assets. The choice of signal distributions is constrained by the investor's information capacity, a constraint on the total informativeness of the signals he can observe. 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 In order to study information choices, we want to begin by modeling investors who benefit from acquiring information. Therefore, we give investors a preferences for early resolution of uncertainty. Investors, with absolute risk aversion parameter ρ, maximize their expected certainty equivalent wealth:

$$U = E_1 \left\{ -\log \left(E_2[exp(-\rho W)] \right) \right\}. \tag{1}$$

Utility can instead be defined over consumption by assuming that all wealth is consumed at the end of period 3. The term $-\log \left(E_2[exp(-\rho W)] \right)$ is the level of consumption that makes the investor indifferent between consuming that amount for certain and investing in his optimal portfolio, in period 2. This certainty equivalent consumption is conditional on the realization of the signals the investor has chosen to see. Since these signals are not known in period 1, the investor maximizes the expected period-2 certainty equivalent, conditioning on information in prior beliefs.

This formulation of utility has the desirable feature that it treats learned information and prior information as equivalent. It does so without losing the exponential structure of preferences that will keep the problem tractable.

Budget Constraint Let $r > 1$ be the risk-free return and q and p be Nx1 vectors of the number of shares the investor chooses to hold and the asset prices. Investor's terminal wealth is then his initial wealth W_0, plus the profit he earns on his portfolio investments:

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

Initial information We model two countries, home and foreign. Each has an equal-sized continuum of investors, whose preferences are identical. Home and foreign 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. We will call this difference in variances a group's information advantage.

Information acquisition At time 1, investors choose how much to learn about each asset's payoff. This choice is equivalent to choosing the variance-covariance matrix Σ_η of a normally-distributed N-dimensional signal η about asset payoffs.[1] Each investor gets a signal drawn from his chosen distribution that is independent of the signals drawn by other investors. The independence assumption is not crucial, but makes aggregation easier.

When asset payoffs co-vary, learning 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

[1]In principle, investors can choose the kind of distribution from which they want to draw signals as well. In this setting, normally distributed signals are optimal. When an objective is quadratic, normal distributions maximize the entropy over all distributions with a given variance (see Cover and Thomas, 1991, chapter 10). Our objective will turn out to have a quadratic form.

factors and the risk of each factor. Learning is then a choice of how much to reduce the variance of each independent risk 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 by 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 risk from the same set of risk factors, but each starts out knowing a different amount about each risk factor.

Without loss of generality, we bypass the choice of signals and model the choice over the posterior beliefs directly. Since sums, products and inverses of prior and signal variance matrices will all have eigenvectors Γ, posterior beliefs will have the same set of risk factors. We denote posterior beliefs with a hat. We can express posterior variance $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$, where Γ is taken as given and the diagonal eigenvalue matrix $\hat{\Lambda}$ is the choice variable. In other words, holding the composition of the risk factors they face constant, investors choose how much to reduce the risk of each factor. The decrease in risk factor i's variance $(\Lambda_i - \hat{\Lambda}_i)$ captures how much an investor learned about that risk.

Learning about risk factor i's payoff $(f'\Gamma_i)$ means that this investor is learning about the ith principal component of asset payoffs. Nothing prevents the investor from learning about many of these principal components. The only thing this rules out is seeing a signal that contains correlated information about risks that are independent. Learning about risk factors (principal components analysis) has long been used in financial research as well as among practitioners. It approximates the kinds of risk categories that investors might consider: business cycle risk, industry-specific risk, firm-specific risk, etc. This paper's main result, that investors learn about risks that they have an initial advantage in, relies on gains to specialization and strategic substitutability in learning; neither force depends on this assumption. However, this risk factor structure makes describing and aggregating information choices tractable.

There are 2 constraints governing how the investor can choose his signals about risk factors. The first is the *capacity constraint*; it governs the quantity of information the investors is allowed to observe. The work on information acquisition with one risky asset quantified information as the ratio of variances of prior and posterior beliefs (Verrecchia, 1982). We generalize the metric to a multi-signal setting by bounding the ratio of the generalized prior variance to the generalized posterior variance, $|\widehat{\Sigma}| \geq e^{-2K}|\Sigma|$, where generalized variance refers to the determinant of the variance-covariance matrix. Capacity K can then be interpreted as the percentage by which an investor can decrease the risk he faces, where risk is measured as the generalized standard deviation of asset payoffs. We assume that K is the same for all investors (section 4.2 relaxes this assumption). This capacity constraint is one possible description of a learning technology. We think it is a relevant constraint because it is a commonly-used distance measure in econometrics (a log likelihood ratio) and in statistics (a Kullback-Liebler distance); it is equivalent to a bound on entropy reduction, which has a long history in information theory as a quantity measure for information (Shannon 1948); it can be re-interpreted as a technology for reducing measurement error; it is a measure of information complexity (Cover and Thomas 1991), and it has been used to describe limited information processing ability in economic settings by Sims (1998). Since determinants are the product of eigenvalues, the capacity constraint is

$$\prod_i \hat{\Lambda}_i \geq e^{-2K} \prod_i \Lambda_i. \tag{3}$$

This particular technology is a strategically neutral and tractable way to describe a rich choice set of signals. It differs from the technology in Mondria (2005) because it requires investors to use capacity to infer asset-payoff relevant information from prices. We endow investors with K large enough to process price information. This assumption prevents there being strategic motives for learning introduced solely by the technology. It requires the same amount of capacity to observe a given signal, whether others observe that signal or not. This constraint also implies that with infinite capacity, all risk is learnable. In Van Nieuwerburgh and Veldkamp (2005), we relax this assumption. This introduces decreasing returns to learning about one risk and makes the specialization result less extreme. But it does not change the conclusion that investors prefer to learn about what they already have an advantage in. Similarly, endogenizing the choice of how much capacity to acquire would not change the decision of how to allocate that capacity, as long as cost was any increasing function of the reduction in generalized variance.

The second constraint is the *no negative learning constraint*: the investor cannot acquire signals that transmit negative information. Without this constraint, the investor might choose to increase uncertainty about some risks so that he could decrease uncertainty further in other variables without violating the capacity constraint. Since negative learning, or intentional forgetting, does not make sense in this context, we rule this out by requiring the variance-covariance matrix of the signal vector η, Σ_η, to be positive semi-definite. Since a matrix is positive semi-definite when all its eigenvalues are positive, the constraint is:

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

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.

Since prices are equilibrium objects, the information they contain depends on the solution to the model. For now, we conjecture that prices are linear functions of the true asset payoffs such that $(rp - A) \sim N(f, \Sigma_p)$, for some constant A. This conjecture is verified in proposition 2.

An investor j's posterior belief about the asset payoff f, conditional on a prior belief μ^j, signal $\eta^j \sim N(f, \Sigma^j_\eta)$, and prices, is formed using Bayesian updating:

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

with variance that is a harmonic mean of the signal variances:

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

These are the conditional mean and variance that investors use to form their portfolios in (10).

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 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 clearing condition is

$$\int_0^1 (\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr)dj = \bar{x} + x. \tag{7}$$

Definition of Equilibrium An equilibrium is a set of asset demands, asset prices and information choices, such that

1. Given prior information about asset payoffs $f \sim N(\mu, \Sigma)$, each investor's information choice $\hat{\Lambda}$ maximizes (1), subject to the capacity constraint (3) and the no-negative-learning constraint (4);

2. Given posterior beliefs about asset payoffs $f \sim N(\hat{\mu}, \widehat{\Sigma})$, each investor's portfolio choice q maximizes (1), subject to the budget constraint (2);

3. Asset prices are set such that the asset market clears: (7) holds;

4. Beliefs are updated, using Bayes' law: (5) and (6);

5. Rational expectations hold: period-1 beliefs about the portfolio q are consistent with the true distribution of the optimal q.

We rewrite period-2 expected utility to eliminate the period-2 expectation operator. In period 2, the only random variable is $f \sim N(\hat{\mu}, \widehat{\Sigma})$. Using the formula for a mean of a log normal, substituting in the budget constraint (2), and substituting $\Gamma\hat{\Lambda}\Gamma$ for $\widehat{\Sigma}$, we can restate the optimal learning and investment problem as choosing portfolios and posterior risk factor variances to maximize the expectation of a standard mean-variance objective:

$$\max_{q,\hat{\Lambda}} E\left[\rho q'(\hat{\mu} - rp) - \frac{\rho^2}{2} q'\Gamma\hat{\Lambda}\Gamma' q | \mu, \Sigma\right]. \tag{8}$$

subject to (3) and (4)

2 Why Might Information Advantages Disappear?

Taken at face value, theories that explain the home bias by relying on an initial information advantage seem unappealing. The problem with assuming that informational advantages will automatically lead to a home bias is illustrated in the context of a model where investors choose what to learn, in order to minimize the variance of a given portfolio. In this setting, an investor who starts out with more information about one asset will undo that advantage by learning about every other asset, until he runs out of capacity, or is equally uncertain about all assets. As Karen Lewis (1999) puts 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."

2.1 A Model without Increasing Returns to Information

Rather than regarding the portfolio as an endogenous choice variable, suppose the investor takes q as given when choosing what to learn. It is this assumption that shuts down the increasing returns to scale.

Let $\tilde{q}_i = \Gamma_i' q$. This represents the amount of risk factor i that an investor holds in his portfolio. Then the objective (8) collapses to choosing $\hat{\Lambda}_i$'s to minimize $\sum_i \tilde{q}_i^2 \hat{\Lambda}_i$, subject to the product constraint (3) and the no-forgetting constraint $\Lambda_i - \hat{\Lambda}_i \geq 0 \ \forall i$. The first-order condition of the Lagrangian problem describes the optimal learning rule.

Proposition 1 ***Learning Undoes Information Advantages*** *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 in appendix A. What is important about this result is that the investor has a target posterior variance for each risk ($\frac{1}{\tilde{q}_i^2} M$), that does not depend on prior variance. An initial information advantage in one risk factor may cause the prior variance to be less than the target posterior variance. In this case, the investor would like to forget some of the information he knows, in order to bring his posterior variance up to his target. The no-negative learning constraint keeps him from forgetting. Instead, the investor chooses not to learn any more about this risk and devotes all his capacity to learning about other risks whose variances are still above their target levels. Learning about the most uncertain risks undoes an investor's information advantage.

If the investor has sufficient capacity, he can fully compensate for any initial information advantage he was given. If this is the case, then no matter what the investor has local knowledge of, he will always end up with the same posterior beliefs after learning.

Corollary 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 $k \neq i$.*

Proof in appendix A.1.

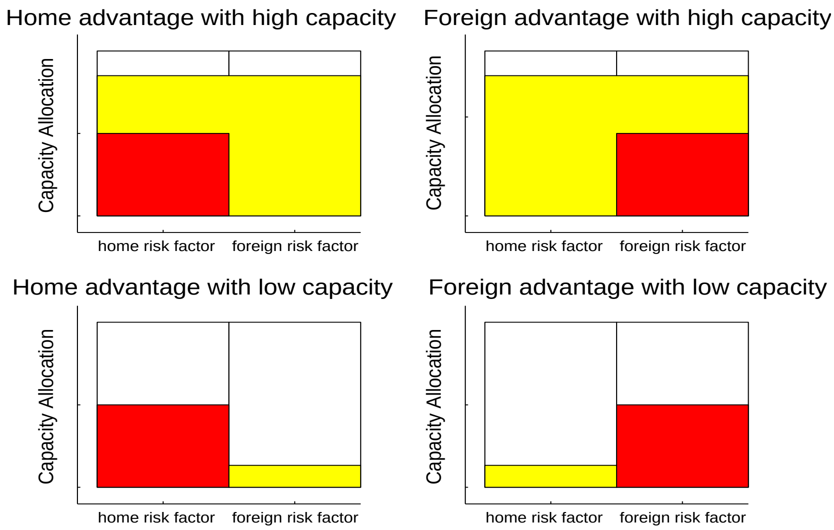

Figure 1: Allocation of information capacity for a low and high-capacity representative investor. The lightly shaded area represents the amount of capacity allocated to the factor. The dark area represents the size of the information advantage. The unfilled part of each bin represents the posterior variance of the risk factor $\hat{\Lambda}_i$. With high capacity, adding the dark block to either bin would result in the 'water level' $\hat{\Lambda}$ 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 graphically. 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). We illustrate a case where there is a home and foreign risk factor and $\tilde{q}^{home} = \tilde{q}^{foreign}$; the two bins are equally deep because both risk factors are equally valuable to learn about. 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. Learning choices compensate for initial information advantage in such a way as to render the nature of the initial advantage irrelevant. Having an initial advantage in home risk will result in the same the same posterior variances for home and foreign assets as having an advantage in foreign risk. Since the asset holdings depend on posterior variances, the allocation to home and foreign assets is the same. With sufficient capacity, initial information advantages cannot contribute to a home bias.

The bottom two panels of figure 1 illustrate capacity allocation when capacity is low. The investor would like to have the water level (his target posterior precision) be the same in both bins. The no-forgetting constraint prevents him from breaking up the brick to achieve an equal water level in both bins; he cannot equalize uncertainty across risk factors. The constrained optimal solution is for the investor to devote all his capacity to learning about the risk factor he is most uncertain about.

2.2 Mechanisms to Preserve Information Advantages

Without increasing returns, there are two ways that initial information advantages can persist: low capacity or unequal processing costs. When capacity is low relative to the initial advantage (as in the bottom panels of figure 1), more precise posterior beliefs for home assets generates a home bias. However, if this explanation were true, then individuals would never choose to learn about local 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. Furthermore, Pastor (2000) shows that even an investor, with no capacity to acquire signals, who passively observes all return realizations, must have implausibly precise prior beliefs to justify the observed home bias.

The second candidate explanation is that investors have a harder time processing information about foreign assets. We investigate a simple 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.[2] We replace (3) with a capacity constraint that requires ψ times more capacity to process foreign than home information:

$$\frac{1}{2}\left[\log(\sigma_h) - \log(\hat{\sigma}_h)\right] + \frac{\psi}{2}\left[\log(\sigma_f) - \log(\hat{\sigma}_f)\right] \leq K. \tag{9}$$

Next, we look at the optimal learning choice and the resulting optimal portfolio. Taking first order conditions with respect to $\hat{\sigma}_h^2$ and $\hat{\sigma}_f^2$ and rearranging yields: $\hat{\sigma}_f^2/\hat{\sigma}_h^2 \leq \psi q_h^2/q_f^2$. Capacity permitting, an investor will set the ratio of posterior variances to $\psi q_h^2/q_f^2$. Thus, for an investor

[2] When home and foreign assets are correlated, it is difficult to disentangle whether a given piece of information is home or foreign. The assumption of zero correlation between home and foreign assets has two effects on this ψ estimate. First, it will make the gains to diversification large and overestimate the benefits of learning about foreign assets. This will bias ψ upward. Second, if home signals are partially informative about correlated foreign assets, home bias would be lower. As a result, the friction ψ would have to be higher to explain the large home bias.

that initially expects to hold a balanced home-foreign portfolio ($q_h = q_f$), $\psi \geq \hat{\sigma}_f^2/\hat{\sigma}_h^2$. Having chosen what to learn and observed the chosen signal, the optimal portfolio for the investor with exponential utility is: $q^\star = \frac{1}{\rho}\widehat{\Sigma}^{-1}(\hat{\mu} - pr)$. This portfolio will generally not be what the investor expected to hold in period 1 ($q \neq q^\star$). If home and foreign assets have the same expected return ($\hat{\mu} - pr$), then $\frac{q_h^\star}{q_f^\star} = \frac{(\hat{\sigma}_h^2)^{-1}}{(\hat{\sigma}_f^2)^{-1}} = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_h^2}$. Since the average U.S. investor holds 7.3 times more home assets than foreign assets, ψ must be at least 7.3 to explain home bias. 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 still arise (and required processing costs would fall) if an investor anticipated holding a lot of the home assets: $q_h > q_f$. But then home bias would arise not from processing costs, but from portfolio expectations. This is exactly the mechanism explored in section 3.

Is this cost ratio ψ realistic? The model's predicted relative *shadow* price of foreign information ($\psi = 7.3$) seems out of line with various measures of the *market* price of foreign information. First, English versions of financial newspapers from Germany, France, Spain, Italy and the UK are inexpensive and easy to access. Second, average salaries for translators are typically 25% less than for financial analysts.[3] If producing home information required one analyst, and foreign information required one analyst and one translator, then the translator's salary would have to be 6.3 times the analyst's. Third, translating a 5000-page report costs approximately \$900.[4] If $\psi = 7.3$, a 5000-page domestic research report must cost no more than \$150. It is possible that agency problems and legal/accounting differences add information costs, but the size of the costs must be large.[5]

The model discussed in this section shut off the increasing returns to information mechanism, by holding investors' portfolios fixed when they choose what to learn. Sustaining information asymmetry is an uphill battle.

[3] Average salary figures from PayScale.com for New York state. In other states such as Illinois, Florida and Texas, translators are paid only 40-60% of the salary of financial analysts.

[4] Source: Click2Translate.com cost estimate for translation by a native speaking translator from German to English.

[5] Importantly, many costs associated with learning about foreign assets, such as understanding the legal environment and the tax treatment of foreign earned income, tend to be fixed costs. Fixed costs cannot explain the observed lack of diversification, because after they are sunk, the investor should invest in a well diversified foreign portfolio. Kang and Stultz (1997) and Seasholes (2004) find that foreign investment tends to be concentrated in a country's large, high-beta assets. This only makes sense if there are benefits to specialization (see section 4.2).

3 A Rational Expectations Model of Specialized Learning

This section analyzes a model where small differences in investors' information not only persist, but are magnified by the increasing returns to learning. The only change in the model 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 the assumption of information immobility is a defensible one. It is not that home investors can't learn foreign information; they choose not to. They make more profit from specializing in what they already know.

3.1 The Period-2 Portfolio Problem

We solve the model using backwards 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

$$q^j = \frac{1}{\rho}(\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr). \tag{10}$$

Aggregating these asset demand across investors and imposing the market clearing condition (7) delivers a solution for the equilibrium asset price level.

Proposition 2 *Asset prices are a linear function of the asset payoff and the unexpected component of asset supply:* $p = \frac{1}{r}(A + f + Cx)$.

Proof is in appendix A.2, along with the formulas for A and C.

3.2 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 (10), into expected utility (8). Combining terms yields

$$U = E\left[\frac{1}{2}(\hat{\mu}^j - pr)'(\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr)|\mu, \Sigma\right]. \tag{11}$$

At time 1, $(\hat{\mu}^j - pr)$ is a normal variable, with mean $(-A)$ and variance $\Sigma_p - \widehat{\Sigma}^j$.[6] Thus, expected utility is the mean of a chi-square. Using the fact that the choice variable $\hat{\Lambda}$ is a diagonal matrix, that $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$, that $-\Gamma_i' A$ is as in equation (18), and the formula for the mean of a chi-square, we can rewrite the period-1 objective as:

$$\max_{\hat{\Lambda}^j} \sum_i \left(\Lambda_{pi} + (\rho\Gamma_i'\bar{x}\hat{\Lambda}_i^a)^2\right)(\hat{\Lambda}_i^j)^{-1} \quad s.t.\ (3) \text{ and } (4) \tag{12}$$

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.

3.3 Results: Learning with Increasing Returns

The key feature of the learning problem (12) is that it is convex in the posterior variance. It is the convexity of the objective that delivers us a corner solution. The solution is to reduce variance on one risk factor as much as possible.

Proposition 3 ***Optimal Information Acquisition*** *In general equilibrium with a continuum of investors, each investor j's optimal information portfolio 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:*** $\frac{\hat{\Lambda}_i^a}{\Lambda_i^j}\rho^2(\Gamma_i'\bar{x})^2\hat{\Lambda}_i^a + \frac{\Lambda_{pi}}{\Lambda_i^j}$.

Proof: See appendix A.3.

Three features make a risk factor 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 A.2.) Third, and most importantly for the point of the paper, the investor should learn about risk factors that he had 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 expects to hold, based on his prior information, is the factor's

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

expected return, divided by its variance: $\Lambda_i^{-1}\rho\hat{\Lambda}_i^a\Gamma_i'\bar{x}$. 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 decreases its expected posterior variance $\hat{\Lambda}_i$. Re-computing the expected portfolio with variance $\hat{\Lambda}$, instead of Λ, further increases 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, is what generates the increasing returns to specialization.

Aggregate Learning Patterns Learning is a strategic substitute. 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. Consider constructing this Nash equilibrium by an iterative choice process. The first investor will begin by learning about the risk with the highest learning index. Suppose there is another risk factor j whose 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 j. Ex-ante identical investors will learn about different risks. All home investors 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 foreign risks that are learned 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 will depend on the country-wide capacity. The within-country equilibrium capacity allocation is described in Van Nieuwerburgh and Veldkamp (2005). Despite the fact that many risk factors are potentially being learned about in equilibrium, it remains true that each investor learns about one of these factors.

Learning and Information Asymmetry 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 $\star$ superscript on each of these matrices denotes foreign belief counterparts. Then, for example, $\log(|\Lambda_f|)$ represents home investors' prior uncertainty (entropy) about foreign risk factors and $\log(|\hat{\Lambda}_h^\star|)$ represents foreigners' posterior uncertainty about home risks.

Corollary 2 ***Learning Amplifies Information Asymmetry: symmetric markets*** *If for* *every home factor hi, there is a foreign factor fi such that $\Lambda_{hi} = \Lambda^{\star}_{fi}$ and $\Gamma_{hi}\bar{x} = \Gamma_{fi}\bar{x}$, then home investors will learn exclusively about home risks and foreign investors will learn exclusively about foreign risks.*

Proof: See appendix A.4.

When risk factors are symmetric, an investor with no information advantage would be indifferent between learning about home and foreign risks. A slight advantage in home risk delivers a strict preference for specializing in that risk. This effect can be seen in the learning index: an information advantage in risk i implies that the variance of prior beliefs Λ_i^j is low. A low Λ_i^j increases the value of the learning index and makes learning about risk i more desirable. Since investors with no information advantage are indifferent, *any size initial advantage* tilts preferences toward learning more about home risks and amplifies the initial advantage.

Corollary 3 ***Learning Amplifies Information Asymmetry: general case*** *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 $\frac{|\hat{\Lambda}_f|}{|\hat{\Lambda}^{\star}_f|} \geq \frac{|\Lambda_f|}{|\Lambda^{\star}_f|}$.*

Proof: See appendix A.4.

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. Likewise, foreigners have a higher index 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 like the principle of comparative advantage: If country A has an advantage in apples and country B an advantage in bananas, the only 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.

The more asymmetric the markets, the less learning will amplify information asymmetry. In the most extreme asymmetric case, the initial advantage will just be preserved. For example, if the home market is much smaller than foreign, then all investors might 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.

3.4 Home Bias in Investors' Portfolios

To explore the implications of the theory for home bias, we first need to define a benchmark diversified portfolio. We consider two benchmarks. The first portfolio is one with no information advantage and no capacity to learn. If home investors and foreign investors have identical posterior beliefs, they hold identical portfolios. Actual portfolios depend on the realization of the asset supply shock. The expected portfolio as of time 1 for each investor is equal to the per capita expected supply $\bar{x}$.

$$E[q^{no\ adv}] = \bar{x} \tag{13}$$

A second natural benchmark portfolio is one where investors have initial information advantages, but no capacity ($K = 0$) to acquire signals and do not learn through prices. This is the kind of information advantage that Ahearne, Griever and Warnock (2004) capture when they estimate the home bias that uncertainty about foreign accounting standards could generate.

$$E[q^{no\ learn}] = \Gamma\Lambda^{-1}\Lambda^{a}\Gamma'\bar{x}, \tag{14}$$

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 ($\bar{x}$ in equation (15)) 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 an initial information advantage and capacity to learn $K > 0$ is:

$$E[q] = \Gamma\hat{\Lambda}^{-1}\hat{\Lambda}^{a}\Gamma'\bar{x} \tag{15}$$

Learning has two effects on an investors portfolio. The first is that it magnifies the position he decides to take, and the second is that it tilts the portfolio towards the assets learned about. The magnitude effect can be seen from equation (10). The information advantage, coupled with learning choices, reduces home investors' risk of investing in home assets. ($\hat{\Sigma}_i^{-1} = \Gamma\hat{\Lambda}_i^{-1}$ is high for home risks i.) Lower risk makes investors want to take larger positions, positive or negative, in the asset. But why should the position in home assets be a large long position, rather than a large short one? The direction effect, which is a general equilibrium effect, makes home investors want to hold a *positive* quantity of their home assets. 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, pushes up the asset's return. Home investors are being compensated for more risk than they bear ($\hat{\Lambda}_i^a > \hat{\Lambda}_i^j$ in equation 15). Based on their information, this asset delivers high risk-adjusted returns. High returns make a long position optimal, on average. It is still possible that a very negative signal realization would make home investors want to short home assets, but the expected portfolio holding is long. Both the information advantage and the general equilibrium effect increase home bias as capacity rises.

The next two propositions formalize the difference between the optimal portfolio (15), and the benchmark portfolios (14) and (13). 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.

Proposition 4 ***Information Mobility Increases Home Bias: symmetric markets*** *If for every home factor hi, there is a foreign factor fi such that $\Lambda_{hi} = \Lambda_{fi}^{\star}$ and $\Gamma_{hi}\bar{x} = \Gamma_{fi}\bar{x}$, then every home investor's expected portfolio contains more of assets that load on home risk when he can learn ($K > 0$), than when he cannot ($K = 0$): $\Gamma_h' E[q] > \Gamma_h' E[q^{no\ learn}] > \Gamma_h' E[q^{no\ adv}]$.*

Proof: See appendix A.5.

When asset markets are symmetric, every investor learns exclusively about their home risk factors (corollary 2). Because of the information and general equilibrium effects, learning (information mobility) increases the expected home asset position.

The extent of the home bias depends on what this investor knows relative to the average investor. When $K = 0$, the posterior variance $\hat{\Lambda}$ is the same as the prior variance Λ, and equal to the average variance ($\hat{\Lambda}^a = \Lambda^a$). The optimal portfolio is the no-learning portfolio ($q = q^{no\ learn}$). As capacity K rises, the posterior variance falls on the assets the investor learns about ($\hat{\Lambda}_i^{-1}$ rises), and those assets become more heavily weighted in the portfolio. The more capacity an investor has, the more their portfolio is tilted away from the diversified portfolio and towards the assets they learn about.

Proposition 5 ***Information Mobility Increases Home Bias: general case*** *The average home investor's portfolio contains at least as much of assets that load on home risk when he can learn ($K > 0$), than when he cannot ($K = 0$):* $\Gamma_h' E[q] \geq \Gamma_h' E[q^{no\ learn}] > \Gamma_h' E[q^{no\ adv}]$.

Proof: See appendix A.5.

When market size is different across countries, corollary 3 shows that no home investor will learn more about foreign risks than any foreign investor will, and vice versa. Therefore, the average home investor knows more about home risks, and tilts his portfolio to hold more of them.

In the most extreme case, all investors learn about one risk (for example because that risk is large). The ratio of home investors' and the average investor's posterior variance is the same whether investors have positive or zero capacity. This implies that $E[q] = E[q^{no\ learn}]$. In this extreme case, learning does not amplify the home bias, but it doesn't undo it either as in section 2 ($E[q] > \bar{x}$).

When home risk factors are small, home investors are more likely to learn about larger foreign risks, and reduce the average level of home bias. This prediction fits with evidence on the cross-country patterns of home bias. Small risk-factor countries should have smaller financial markets and assets with lower world-market betas. Small countries such as Belgium, The Netherlands, and Scandinavian countries all have less home bias than the U.S., Japan or larger European countries (Morse and Shive, 2003).

The next proposition shows that home investors earn higher returns on home assets. Following Admati (1985), we define the excess return on asset i as $(f_i - p_i r)$.

Proposition 6 ***Better-Informed Investors Earn Higher Returns*** *As capacity K rises, the expected return an investor earns on the component of his portfolio that he learns about, rises:* $\partial E[(\Gamma_i' q)(\Gamma_i'(f - pr))]/\partial K > 0$.

Proof: See appendix A.6. Home investors earn excess returns on home assets. This is consistent with evidence found by Hau (2001). Foreigners don't learn about home assets, but hold them as part of their diversified portfolio. The home investor profits from his superior information on home assets. The more learning capacity the home investor has, the stronger the information advantage.

It is not the case that an investor raises his expected profit $(f_i - p_i r)$ on any one share of asset i by learning. Payoffs f are exogenous and prices p are determined by the average investor. Rather, a better-informed investor takes a larger position in the assets that he learns about, and increases the correlation of his asset demand q with asset payoffs $(f - pr)$.[7] He holds a long position in the asset when it is likely to pay high returns, and shorts the asset when it is likely to pay very low returns.

3.5 A Numerical Example

In this section we illustrate the magnitude of the home bias that our learning mechanism can generate, by way of a numerical example. We start from the symmetric setup outlined in proposition 3. There are two countries, home and foreign, with a large number of investors (1000) in each. The risk aversion parameter of all investors is $\rho = 2$. There are N assets in the economy, $N_h = 5$ home assets and $N_f = 5$ foreign assets. We consider a case with symmetric information advantages: home investors face proportionately less risk from the first $N/2$ factors and foreign investors have an identical risk reduction for the last $N/2$ factors. Then prior variance eigenvalues are $(1+\alpha)\Lambda_h = \Lambda_h^\star$ and $\Lambda_f = (1+\alpha)\Lambda_f^\star$. The parameter $\alpha \geq 0$ measures the strength of the information advantage.

Uncorrelated assets In the first exercise, all assets in the economy are mutually uncorrelated. The eigenvector matrix Γ is the identity matrix. The supply of each asset is set to one ($\bar{x}$ =1), making the quantity of each risk factor one as well ($\Gamma'\bar{x} = 1$). The asset supply variance is set to make prices as informative as signals. Expected payoffs for home and foreign assets are equal. They are equally spaced between 1-2% per period (per month). The mean of the average investor's prior belief is equal to the true asset payoffs. The standard deviation of prior beliefs is first set between 5-10%, such that all assets have the same expected payoff to standard deviation ratio. Then, the variance of home investors' prior beliefs about home assets is set ten percent lower than foreign investors' variance ($\alpha = 0.1$). The reverse is true for foreign assets. We vary learning capacity K to explore its effect.

[7] $\text{Corr}(q, f - pr) = \text{corr}(\hat{\mu} - pr, f - pr)$. Since, $\hat{\mu}$ is an unbiased expectation of f, it is equal to $f + \epsilon$, where ϵ is orthogonal to f and pr. The variance of ϵ is the variance of f, conditional on $\hat{\mu}$, which is $\widehat{\Sigma}$. Thus, $\text{corr}(q, f - pr) = (var(f - pr) + \widehat{\Sigma})^{-1/2} var(f - pr) var(f - pr)^{-1/2}] = (var(f - pr) + \widehat{\Sigma})^{-1/2} std(f - pr)$. As capacity increases, expectation error variance $\widehat{\Sigma}$ falls, and correlation rises.

Following convention, we define the home bias as

$$\text{home bias} = 1 - \frac{1 - share\ of\ home\ asset\ in\ home\ portfolio}{share\ of\ foreign\ assets\ in\ world\ portfolio} \tag{16}$$

In this example, the share of foreign assets in the world portfolio is 0.5. The share of foreign assets in the home portfolio is an average over all home investor portfolios.

The first benchmark is a world where there is no initial information advantage and no learning capacity. The home bias is zero. The second benchmark is an economy with initial information advantage ($\alpha = 0.1$), but no learning capacity (K=0). This calculation represents the home bias that is purely due to the initial advantage. The ten percent initial information advantage leads to a 5.3 percent home bias in asset holdings. Next, we switch on the learning mechanism by giving each investor a positive capacity K. Learning substantially magnifies the home bias because each home investors specialize in learning about one home asset. The magnification increases steeply in the learning capacity. When there is only enough capacity to view a signal that eliminates 20 percent of the risk in one asset (K=.22, $1 - e^{-K} = .20$), the home bias almost doubles from 5.3 to 9.3 percent. When there is enough capacity to eliminate 61 percent of the risk in one asset (K=.95, $1 - e^{-K} = .61$), the home bias is 31.4 percent, six times as large as without learning effect.

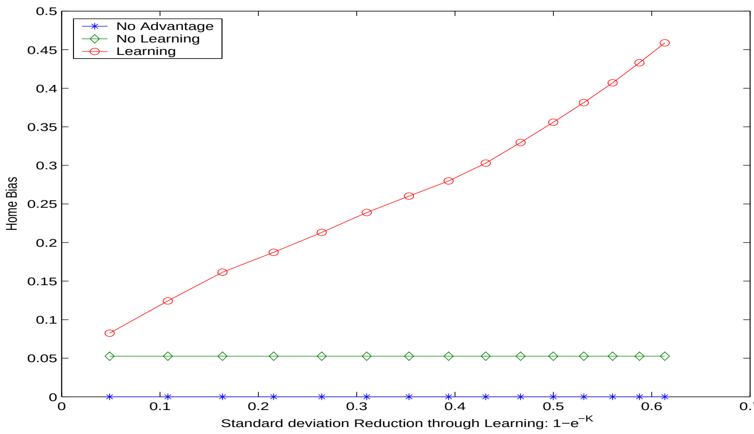

Figure 2: Home Bias Increases With Capacity.
Numerical example with two countries. Assets within a country have correlated payoffs (cov= $.03^2$). Home bias is defined in (16). The 'no advantage' line (stars) gives the home bias in an economy with no initial informational advantage and no capacity to learn. The 'no learning' line (diamonds) refers to the home bias in a world with a small initial information advantage (10%) and no learning capacity. The 'learning' line (circles) plots the home bias in our model. The initial information advantage is 10% and the learning capacity K varies between 0.05 and 0.85. The horizontal axis plots the potential percentage reduction in the standard deviation of one asset, $1 - e^{-K}$.

Correlated Home Assets When home assets are positively correlated with each other (covariance of $.03^2$), and foreign assets are positively correlated with each other (covariance of $.03^2$), but the two sets of assets are mutually uncorrelated, each investor learns about one risk factor, that all his domestic assets load on. Introducing correlation doubles the home bias: Home bias is 18.7% when $1 - e^{-K} = .20$ and increases to 45.9% for $1 - e^{-K} = .61$. (See line with circles in figure 2.) In contrast, the no learning benchmark is virtually unaffected (5.3%, line with diamonds). The reason for this increase is that the set of assets an investor learns about becomes more diversified. Such portfolio is less risky to hold. Therefore, the investor takes a larger position and holds more of the home risk factor. The home bias in the numerical example does not match the 76% home bias in the U.S. portfolio data, for the levels of capacity we explored. In the next section, we endow investors with advantages in local risks, rather than country risks. This change in the information structure allows the model to match the data.

4 Extensions

4.1 Local Investing: Heterogeneous Home Information

In the model of section 3, all home investors have the same precision signals. Because they were ex-ante identical, home investors were indifferent between holding their portfolio or the portfolio held by any other home investor. Coval and Moskowitz (1999) have shown that many investors prefer not only home assets, but local assets. By giving investors slightly more precise signals over local assets, this model can explain the local investment bias, and the accompanying excess returns from investing locally.

Suppose that home investors each had an advantage in only one home risk factor, the one most concentrated in their region's asset. We assume, without loss, that there is one asset per region. An investor j from region m draws an independent prior belief from the distribution $\mu^j \sim N(f, \Sigma_m)$, 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. With this setup, all of the results of section 3 still hold.

Investors from various localities will have an incentive to learn more about their local assets, because of the pre-existing information advantage (proposition 3). Preserving, or amplifying local information advantages causes expected portfolios to weight local assets more heavily, and to make higher expected profits from trading local assets (proposition 6). This is consistent with evidence that local investments earn an extra 2% risk-adjusted return per year (Coval and Moskowitz, 2001). The assets for which local advantage is most valuable are assets that others are not learning about (high Λ^a). Ivkovic, Sialm and Weisbenner (2004) show that individual investors, who hold portfolios that are concentrated in local stocks, not listed on the S&P 500, make returns that are 7% higher than what they would earn on a diversified portfolio.

A unified explanation for home bias and within-country local bias is something that many theories of home bias cannot provide. Explanations rooted in exchange rate risk, institutional difference, or language barriers do not apply to differences in portfolios across U.S. regions. The fact that a home bias is present both within and across national borders, makes an information-based explanation appealing. An extreme example of specialized information about local assets is information about one's own firm. With a tiny advantage in an investor's own firm, the investor would rationally learn more about his firm and overweight it in his portfolio. However, a detailed analysis of the own-company stock puzzle is beyond the scope of this project.

Numerical Example To quantify the local bias, we use the same numerical example as in section 3.5. There is correlation among home and among foreign assets, but no correlation between home and foreign assets. There are 5 regions at home and 5 regions abroad, each with one local asset. The only difference between this exercise and the one in section 3.5 is that instead of giving 1000 home (foreign) investors a 10% initial information advantage in every home (foreign) asset, we give 200 investors each a 50% information advantage in one asset; the aggregate information advantages at home and abroad are unchanged. We measure local bias as:

$$\text{local bias} = 1 - \frac{1 - share\ of\ local\ asset\ in\ local\ portfolio}{share\ of\ non-local\ assets\ in\ world\ portfolio} \tag{17}$$

We could also interpret this as the home bias in a 10-country world. It could also represent a bias toward industries that investors have prior knowledge of, perhaps because of their job.

Without learning, the average local bias is only 5.0%. With learning capacity, $1 - e^{-K} = 0.61$, the average local bias is 35%. This average understates the local bias in many regions; the highest is 49%. A local bias of 35% (49%) implies that the local investor holds 41% (54%) of their portfolio in the local asset. This is 4.1 (5.4) times larger than what full diversification would predict.

We compute the *home bias* in this economy by averaging positions in all home assets across home investors. Without learning, the home bias is 7.5%. Home bias increases with learning, to 22% for low capacity ($1 - e^{-K} = 0.20$) and 72% for high capacity ($1 - e^{-K} = 0.58$), which is close to the 76% home bias observed in the data.

Giving investors a local advantage as opposed to a home advantage in all home assets (section 2.5) generates 26% more home bias (for high capacity), even though the aggregate information advantage is the same. This is because an investor's portfolio share in an asset depends on the difference between the risk he bears and the risk he is compensated for. When fewer investors share an advantage in the same risk, they have a larger advantage relative to the average investor, and earn higher returns. Because the returns to specialization are higher, investors diversify their portfolios less.

4.2 High-Capacity Home, Low-Capacity Foreign Investors

Using foreign investment data from Taiwan, Seasholes (2004) finds that foreign investors outperform the Taiwanese market, particularly when foreigners are investing 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." We can ask of our model: If Taiwanese investors have low capacity, will Americans invest in Taiwanese assets? Will they outperform the market? Will American excess returns be concentrated in assets that load heavily on the largest risk factors? The answer to all three questions turns out to be yes.

When American capacity greatly exceeds Taiwanese capacity, then Americans will invest in Taiwan to capture higher returns. To see why, suppose instead that all investors devote their capacity to learning only about their respective home risks. Since Americans have more capacity, they will reduce the average posterior variance for their assets by more: $\hat{\Lambda}^a_{hi} < \hat{\Lambda}^{a\star}_{fi}$, for equally-sized home and foreign risks hi and fi. Recall that expected returns are determined by average posterior variance; when Americans have higher capacity, expected returns for US assets are lower than for Taiwanese assets. There will be some level of capacity difference that will create a great enough difference in returns to induce some Americans to invest in Taiwan. Expecting to hold Taiwanese assets, some Americans will learn about Taiwan. As Americans learn about Taiwan, they depress Taiwanese returns. This does not mean that returns in Taiwan and the U.S. will be equalized. Those Americans who learn about Taiwan will still face more posterior risk than if they had learned about the U.S., because of their initial information disadvantage in Taiwan. Higher returns in Taiwan compensate Americans for the higher posterior risk they bear.

Although U.S. investors face more posterior uncertainty in Taiwan than in the U.S., they can still outperform the average Taiwanese investor, when trading in the Taiwanese assets they research.[8] If Americans have capacity that exceeds Taiwanese capacity by more than the size of their initial information advantage in one risk factor, then Americans can become better informed about that risk than the average Taiwanese investor. By proposition 6, being more informed than the average investor implies than the American investor will out-perform the average investor in assets that load on his researched risk factor (fi).

What are the Taiwanese risk factors that Americans will learn about? Assuming that the average uncertainty ($\hat{\Lambda}^a$), noise in prices (Λ_p) and American uncertainty (Λ) about each Taiwanese risk is identical, then the most valuable risk to reduce is the one with the largest quantity, the highest $\Gamma_i \bar{x}$ (proposition 3). Americans should learn about, hold more of, and profit from the risk factors that the largest assets weight most heavily on.[9] Thus, the model, and Seasholes' data, both predict that high-capacity foreigners trading large assets, with high market covariance, are likely to out-perform the market.

4.3 Portfolio Flows and Gravity Models

Patterns of learning about foreign assets tell us about what patterns of equity flow volume should look like. The large cross-border equity flows observed have proven difficult to reconcile with many theories of home bias (Tesar and Werner 1995). However, Coval (2000) and Brennan and Cao (1987) have shown that information asymmetry across borders can account jointly for the home bias and the high trade volume. This paper extends those explanations by forwarding a theory that predicts when investors in different countries will be most likely to learn different information. Are these predictions about information asymmetry consistent with the patterns of equity flow volume?

Answering this question is complicated by the fact that learning about foreign assets has two opposing effects on gross equity flows between the home and foreign countries. First, learning makes

[8]Americans may also hold Taiwanese assets for diversification purposes, without learning about them. These American investors will under-perform relative to the locals.

[9]If Americans start out knowing a little bit more about the large foreign risk factor than about smaller risks, the effect is reinforced.

foreign assets less risky to home investors. Facing less risk, they take a larger position in foreign assets. When news prompts home investors to trade, their larger position increases the size of their trade. This is a scale effect: It increases total trade, but not the turnover rate (trade per share held). Second, learning moves home investors' beliefs closer to the true foreign payoff, and closer to the beliefs of the foreign investors, on average. The decrease in information asymmetry makes home and foreign investors less willing to trade with each other. This decreases turnover. Only when beliefs differ will one want to sell when the other wants to buy. On net, learning decreases the turnover rate, but it can increase total trade volume.

In the data, patterns of equity flows line up closely with countries' geographical distance from each other. Tesar and Werner (1994) show that turnover rates between 5 OECD countries are higher for countries that are farther away, and are inversely related to portfolio share. Portes, Rey and Oh's (2001) gravity model estimation shows that geographically close markets, with highly correlated returns, generate larger gross equity flows. All three stylized facts: gross flows, turnover rates, and portfolio shares could be explained by learning.

The model predicts the pattern of information implied by the data, because highly correlated markets are efficient to learn about. High correlation of neighboring country returns means that the risk factors that each country's assets load on most heavily are, to large extent, common. Given an information advantage in a home risk factor, home investors also have some advantage in neighboring country assets that load on the same factor. An initial advantage in such a common risk may lead home investors to specialize in it, and learn about the foreign assets that load on it. High correlation makes segments of neighboring countries have information properties similar to home markets. Direct evidence on information flows bolster this prediction. Portes and Rey (2003) find that nearby markets exchange abundant information: they exhibit high telephone traffic and strong evidence of insider trading.

The other important feature of models is that market size increases portfolio flows. This is consistent with our model because of the increasing returns to information. Learning about large risk factors (big $\Gamma_i \bar{x}$) generates more profit because applying a signal to many shares generates more profit than applying it to only a few (proposition 3). If investors learn about the risk factors of large markets, and learning is related to flow volume, then flow volume should be highest for the markets that are most valuable to learn about: large ones.

Gravity model findings fly in the face of standard investment theory because investors trade most with countries whose assets offer little diversification benefit. It is precisely because these assets are poor diversification devices, that it is efficient to learn about them, to hold them, and therefore to trade them prolifically.

5 Conclusions

This paper examined the common assumption that residents of one region have more information about their region's assets than do non-residents. In particular, it poses the question: If investors are restricted in the amount of information they can learn about risky asset payoffs, which assets would they choose to learn about? We show that an investor, who does not account for the effect of learning on his portfolio choice, chooses to study risks he is most uncertain about. He undoes his initial advantage. But, investors with rational expectations reinforce their initial information advantages. Initial information advantages lead the investor to believe that he will hold slightly more of the asset than the average investor will. Knowing that the asset will be a larger part of his portfolio causes the investor to value learning about it more. An initial information advantage in a risk is like a comparative advantage in learning about that risk. By each specializing in learning about what they have a comparative advantage in, investors increase their information asymmetry.

Thus our main message is that information asymmetry assumptions in international finance are defensible, but perhaps not for the reasons originally thought. We do not need to resort to large information frictions; small frictions will suffice because learning will amplify them. With sufficient capacity to learn, small initial information advantages can lead to a home bias of the magnitude observed in the data.

The results characterizing optimal learning strategies can be applied to a wide range of environments to deliver rich cross-sectional predictions. The theory can be interpreted as one of local bias, and predicts the excess returns observed on local investors' portfolios (Coval and Moskowitz 2001). The theory also predicts when investors will choose not to specialize in home assets, and thus predicts patterns of foreign investment, and foreign investment returns. The prediction that foreigners should hold and profit on large assets that are highly correlated with the market is confirmed in empirical work by Seasholes (2004) and Kang and Stultz (1997). Finally, explaining learning choices delivers a model of information flows, a leading explanatory variable in theories of cross-border equity trade (Coval 2000). The prediction that large information and equity flows arise between countries with correlated returns squares with Portes and Rey's (2003) finding that geographic proximity is highly correlated with information proxies and equity trade.

Future work should focus on building a dynamic model of learning and investing. It could teach us more about the patterns of equity flows. It could also pin down prior beliefs, a free parameters in the static model. Since prior beliefs in a dynamic model must arise from posterior beliefs in a prior period, modeling dynamic learning would restrict the admissible set of prior beliefs and give the theory additional predictive power.

Asymmetry in prior beliefs could arise from risky labor income. Baxter and Jermann (1997) argue that countries' labor income growth and equity returns are highly correlated. Although this correlation worsens the home bias puzzle in a standard model, it bolsters our explanation. Labor income provides an initial information advantage that makes learning more about own country, own industry, and own company assets optimal. Massa and Simonov (2004) document that Swedish investors tilt their portfolio towards the industry they work in, and conclude that this is a rational response to being better informed about their own profession.

An important assumption in our model was that every investor must process their own information. If information capacity is costly, rather than fixed, then paying one portfolio manager to learn about each risk is efficient. How might such a setting regenerate a home bias? By its nature, selling information also generates an agency problem. A solution could involve auditing portfolio managers. A manager from the same region, whose initial information resembles the investor's, may require less capacity to audit. In such a setting, portfolio managers who cater to nearby investors would have incentives to learn that mirror their clients' incentives. They would maximize profit by reinforcing their initial information advantage and specializing in home assets. Our theory could be reinterpreted as pertaining to these portfolio managers.

Information asymmetries play a prominent role in international finance. This paper provides tools that can predict where asymmetries are most likely, and what form they will take. It also offers a cautionary word about building theories around assumptions on information sets. Economic agents can choose to acquire information and learn. Ignoring learning incentives when specifying information structures raises questions about the resulting theories. These theories may be analyzing situations that a utility-maximizing agent who can learn would never face.

References

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

[2] Ahearne, Alan; Griever, William and Warnock, Francis. "Information Costs and Home Bias: An Analysis of US Holding of Foreign Equities." *Journal of International Economics*, v.62, 2004, pp.313-336.

[3] Bacchetta, Philippe and van Wincoop, Eric. "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?" Working paper, 2004.

[4] Baxter, Marianne and Jermann, Urban. "The International Diversification Puzzle Is Worse Than You Think" *The American Economic Review*, Vol. 87(1), March, 1997, pp. 170-180.

[5] Boyle, P.; Uppal, Raman and Wang, Tan. "Ambiguity Aversion and the Puzzle of Own-Company Stock in Pension Plans." Working paper, 2003.

[6] Brennan, Michael and Cao, Henry. "International Portfolio Investment Flows." *Journal of Finance*, v.52, 1997, pp.1851-80.

[7] Calvo, Guillermo and Mendoza, Enrique. "Rational Contagion and the Globalization of Securities Markets." *Journal of International Economics.* 2000 vol. 51(1), pages 79-113

[8] Cohen, Lauren. "Loyalty Based Portfolio Choice." Working paper Chicago GSB, 2004.

[9] Coval, Joshua. "International Capital Flows when Investors Have Local Information." Working paper, 2000.

[10] Coval, Joshua and Moskowitz, Tobias. "The Geography of Investment: Informed Trading and Asset Prices." *Journal of Political Economy.* v.109(4), 2001, pp.811-841.

[11] Coval, Joshua and Moskowitz, Tobias. "Home Bias at Home: Local Equity Preference in Domestic Portfolios." *Journal of Finance* v.54(6), 1999, pp.2045-2073.

[12] Cover, Thomas; and Thomas, Joy. *Elements of Information Theory.* John Wiley and Sons, New York, NY, 1991.

[13] Evans, Martin and Lyons, Richard. "A New Micro Model of Exchange Rate Dynamics." Working paper, 2004.

[14] French, Kenneth and Poterba, James. "International Diversification and International Equity Markets." *American Economic Review* vol. 81(2), 1991, pp.222-226.

[15] Hatchondo, Juan Carlos. "Asymmetric Information and the Lack of International Portfolio Diversification." Working paper Univertsity of Rochester, 2004.

[16] Hau, Harald. "Location Matters: An Examination of Trading Profits." *Journal of Finance*, v.56(5), 2001, pp.1959-83.

[17] Hong, Harrison; Kubik Jeffrey and Stein, Jeremy. "The Only Game in Town: Stock-Price Consequences of Local Bias." working paper, 2004.

[18] Huberman, Gur. "Familiarity Breeds Investment." *Review of Financial Studies*, v.14(3), 2001, pp.659-80.

[19] Ivkovic, Zoran; Sialm, Clemens and Weisbenner, Scott. "Portfolio Concentration and the Performance of Individual Investors." Working paper, 2004.

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

[21] Lewis, Karen. "Trying to Explain Home Bias in Equities and Consumption." *Journal of Economic Literature*, v.37, 1999, pp.571-608.

[22] Luo, Yulei. "Consumption Dynamics Under Information Processing Constraints." Working Paper, 2005.

[23] Massa, Massimo and Simonov, Andrei. "Hedging, Familiarity and Portfolio Choice." Working paper, 2004.

[24] Mondria, Jordi. "Financial Contagion and Attention Allocation." Working paper, 2005.

[25] Morse, Adair and Shive, Sophie. "Patriotism in Your Portfolio." Working paper, 2003.

[26] Moscarini, Giuseppe. "Limited Information Capacity As A Source of Inertia." *Journal of Economic Dynamics and Control*, v.28(10), 2004, pp.2003-2035.

[27] Pastor, Lubos. "Portfolio Selection and Asset Pricing Models." *Journal of Finance* 2000, vol.55(1), p.179-223.

[28] Peng, Lin and Xiong, Wei. "Investor Attention, Overconfidence and Category Learning." *Journal of Financial Economics* forthcoming.

[29] Portes, Richard; Rey, Helene and Oh, Yonghyup. "Information and Capital Flows: The Determinants of Transactions in Financial Assets." *European Economic Review.* v.45, 2001, pp.783-796.

[30] Portes, Ricard and Rey, Helene. "The Determinants of Cross-Border Equity Flows." Working paper, 2003.

[31] Radner, Roy and Stiglitz, Joseph. "A Nonconcavity in the Value of Information." In: *Bayesian Models in Economic Theory.* eds. M. Boyer and R.E. Kihlstrom, Elsevier Science Publishers B.V., 1984.

[32] Seasholes, Mark. "Re-examining Information Asymmetries in Emerging Stock Markets." Working paper, 2004.

[33] Shannon, C.E. "A Mathematical Theory of Communication." *Bell System Technology Journal*, 1948, vol.27, pp.379-423 and 623-656.

[34] Sims, Christopher. "Stickiness", *Carnegie-Rochester Series on Public Policy*, v.49(1), 1998, pp.317-356.

[35] Sims, Christopher. "Implications of Rational Inattention." *Journal of Monetary Economics*, v.50(3), 2003, pp.665-690.

[36] Tesar, Linda and Werner, Ingrid. "Home Bias and High Turnover." *Journal of International Money and Finance.* vol. 14(4), 1995, pp.467-92.

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

[38] Tesar, Linda and Werner, Ingrid. "International Equity Transactions and U.S. Portfolio Choice" in Jeffrey Frankel Ed, Internationalization of Equity Markets, NBER Project Report, University of Chicago Press, 1994, 185-216.

[39] Van Nieuwerburgh, Stijn and Veldkamp, Laura. "Information Acquisition and Portfolio Under-Diversification." Working paper, 2005.

[40] Verrecchia, Robert. "Information Acquisition in a Noisy Rational Expectations Economy." *Econometrica*, 1982, v.50(6), pp.1415-1430.

[41] Wilson, Robert. "Informational Economies of Scale." *Bell Journal of Economics*, 1975, v.6, pp.184-195.

A Proof of Proposition 1

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 $\prod_i \hat{\Lambda}_i \geq \prod_i \Lambda_i e^{-2K}$, 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 $M = \upsilon e^{-2K} \prod_i \Lambda_i$. 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 $\hat{\Lambda}_i > \Lambda_i$. □

A.1 Proof of Corollary 1

If $(\Lambda_i - \epsilon)\tilde{q}_i^2 > M$ for $i = argmin_j (\Lambda_j - \epsilon)\tilde{q}_j^2$, then 1 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 $\log(\Lambda_j) - \log(\hat{\Lambda}_j) = \log(\Lambda_j) - \log(\frac{1}{\tilde{q}_j^2} M) = \log(\Lambda_j \tilde{q}_j^2) - \log\left(\min_i \left((\Lambda_i - \epsilon)\tilde{q}_i^2\right)\right)$, the total capacity required is

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

□

A.2 Proof of Proposition 2

From Admati (1985), we know that equilibrium price takes the form $rp = A + Bf + Cx$, where

$$
\begin{aligned}
A &= -\rho \left(\frac{1}{\rho^2 \sigma_x^2} (\Sigma_\eta^a \Sigma_\eta^{a\prime})^{-1} + (\Sigma_\eta^a)^{-1} \right)^{-1} \bar{x}, \\
C &= -\left(\frac{1}{\rho^2 \sigma_x^2} (\Sigma_\eta^a \Sigma_\eta^{a\prime})^{-1} + (\Sigma_\eta^a)^{-1} \right)^{-1} \left(\rho I + \frac{1}{\rho \sigma_x^2} (\Sigma_\eta^a)^{-1\prime} \right).
\end{aligned}
$$

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.

$$
\begin{aligned}
(\Sigma_\eta^a)^{-1} &= \Gamma (\Lambda_\eta^a)^{-1} \Gamma' = \frac{1}{2} \Sigma^{-1} + \frac{1}{2} (\Sigma^\star)^{-1} + \int_j (\Sigma_\eta^j)^{-1} dj, \\
(\Sigma_p)^{-1} &= \Gamma \Lambda_p^{-1} \Gamma' = \frac{1}{\rho^2 \sigma_x^2} (\Sigma_\eta^a \Sigma_\eta^{a\prime})^{-1}, \\
(\widehat{\Sigma}^a)^{-1} &= \Gamma \hat{\Lambda}_a^{-1} \Gamma' = \frac{1}{\rho^2 \sigma_x^2} (\Sigma_\eta^a \Sigma_\eta^{a\prime})^{-1} + (\Sigma_\eta^a)^{-1}
\end{aligned}
$$

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

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

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 $\Gamma'\Gamma = I$. $\Box$

A.3 Proof of Proposition 3

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 $\prod_{i=1}^N y_i \leq e^{2K}$ 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 $\tau = \frac{e^{2K}}{\prod_{j=1}^N \left(1+\frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right)}$. We endow the investor with enough capacity such that he has spare capacity to acquire private signals after devoting capacity to learning from prices: $\prod_{j=1}^N \left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right) < e^{2K}$ and therefore $\tau > 1$. For future reference define the 'spare capacity' of an investor who learns about risk factor i as

$$\tilde{K}_i = K - \frac{1}{2}\sum_{j \neq i} \log\left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right). \tag{19}$$

□

A.4 Proof of Corollaries 2 and 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 $|\hat{\Lambda}_h| = e^{-2\tilde{K}}|\Lambda_h^{-1} + \Lambda_p^{-1}|^{-1}$ and $|\hat{\Lambda}_f| = |\Lambda_f^{-1} + \Lambda_p^{-1}|^{-1}$. Since foreign investors might learn about home risk, but might not: $|\hat{\Lambda}_h^\star| \geq e^{-2\tilde{K}}|\Lambda_h^\star + \Lambda_p^{-1}|^{-1}$ 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.

Symmetric risk factors First, consider a one-agent deviation from the candidate symmetric equilibrium where all home investors learn about home risk factors and all foreign investors learn about foreign risk factors. In particular, suppose all foreign investors besides one were learning about the equal-sized foreign risk factors (i,j such that $\Gamma_i\bar{x} = \Gamma_j\bar{x}$), in the same proportion as home investors are learning about the equal-sized home risk factors. We assumed that $\Lambda_i = \Lambda_j^\star$ for the equal-sized risk factors. Since an equal fraction of investors is learning about i and j, the average signal precision for each risk, and the precision of the price signal will be equal $(\hat{\Lambda}_i^a)^2/\Lambda_i = (\hat{\Lambda}_j^a)^2/\Lambda_j^\star$ and $\Lambda_{pi} = \Lambda_{pj}$. Therefore, home investors get as much value from learning about i as foreign investors get from learning about j:

$$\frac{\Lambda_{pi}}{\Lambda_i} + \frac{(\hat{\Lambda}_i^a)^2}{\Lambda_i}(\Gamma_i'\bar{x})^2 = \frac{\Lambda_{pj}}{\Lambda_j^\star} + \frac{(\hat{\Lambda}_j^a)^2}{\Lambda_j^\star}(\Gamma_j'\bar{x})^2. \tag{22}$$

Combining (20) and (22) tells us that the foreign investor must strictly prefer learning about foreign risk. So, a one-agent deviation from the equilibrium is not optimal.

Next, consider a multiple-agent deviation from the candidate symmetric equilibrium. Suppose that foreign investors learn about the equal-size risk factors, but in different proportions, or that a mass of foreign investors learns about home risks. Note that fewer investors learning about a risk factor increases the $(\hat{\Lambda}_a)^2/\Lambda$ and Λ_p for that factor. For one of the factors i that home investors learn about, there must be fewer investors learning about the same-sized foreign risk factor j such that $\frac{\Lambda_{pi}}{\Lambda_i} + \frac{\hat{\Lambda}_i^a}{\Lambda_i}(\Gamma_i'\bar{x})^2 < \frac{\Lambda_{pj}}{\Lambda_j^\star} + \frac{\hat{\Lambda}_j^a}{\Lambda_j^\star}(\Gamma_j'\bar{x})^2$. This also implies that the foreign investor must strictly prefer learning about j to i, or to any of the other equally valuable home risk factors. The analogous argument can be made showing that home investors always learn about home risks. □

A.5 Proof of Propositions 4 and 5

Symmetric risk factors By corollary 2, 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 $\hat{\Lambda}_i^{-1} = e^{2K}\Lambda_i^{-1}$. When agents can learn ($K > 0$), let ξ_i denote the fraction of home investors that learn about home risk factor i. Then $(\hat{\Lambda}_i^a)^{-1} = \frac{1}{2}\xi_i e^{2K}(\Lambda_i)^{-1} + \frac{1}{2}(1-\xi_i)(\Lambda_i)^{-1} + \frac{1}{2}(\Lambda_i^\star)^{-1}$. 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 e^{2K}. Using equation (15), describing the portfolio with $K > 0$ and equation (14), describing 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 By corollary 3, 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 (15) by Γ_i'. Here, $(\hat{\Lambda}_i^{-1})^{ha} = \xi_i e^{2K}(\Lambda_i)^{-1} + (1-\xi_i)(\Lambda_i)^{-1}$ is the average posterior precision of home investors about risk factor i. The worldwide average precision is $(\hat{\Lambda}_i^a)^{-1} = \frac{1}{2}\xi_i e^{2K}(\Lambda_i)^{-1} + \frac{1}{2}(1-\xi_i)(\Lambda_i)^{-1} + \frac{1}{2}\xi_i^\star e^{2K}(\Lambda_i^\star)^{-1} + \frac{1}{2}(1-\xi_i^\star)(\Lambda_i^\star)^{-1}$.

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 e^{2K}. As a result, the difference between the learning and the no-learning portfolio on risk factor i is strictly positive: $E[q_i] > E[q_i^{no\ learn}]$.

A.6 Proof of Proposition 6

The expected portfolio return on a risk factor can be expressed as the expected portfolio weight times the expected return of the factor, plus their covariance:

$$\text{Expected return } = (\Gamma_i' E[q])(\Gamma_i' E[(f - pr)]) + \Gamma_i' cov(q, f - pr)\Gamma_i. \tag{23}$$

Using equation (15) and the price formula in appendix A.2,it can be shown that $cov(q, f - pr) = \frac{1}{\rho}(\widehat{\Sigma}^{-1}\Sigma - I)$. Canceling out orthogonal eigenvectors, we can rewrite:

$$\text{Expected return } = \frac{1}{\rho}\left(\hat{\Lambda}_i^{-1}(\Gamma_i' A)^2 + \hat{\Lambda}_i^{-1}\Lambda_i - 1\right). \tag{24}$$

Capacity allows the investor to increase his posterior precision for the risk factor he learns about ($\frac{\partial \hat{\Lambda}_i^{-1}}{\partial K} > 0$). One individual's capacity does not affect the aggregate variable A or the given exogenous variables Γ_i or Λ_i. Therefore, capacity increases the expected return. $\square$