INFORMATION IMMOBILITY AND THE HOME BIAS PUZZLE

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

December 17, 2004

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. Information is least mobile when learning capacity and foreign market capitalization are small.

*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 Mark Wright, Dave Backus, and seminar participants at NYU 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. But this seems to replace the assumption of capital immobility with the equally implausible assumption of information immobility. For example, if an American wished to get information about Japanese firms and markets, she could presumably hire a Japanese investment manager. Competition among managers should make the cost relatively low. 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, 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. The reason the investor doesn't want to learn diverse information is that 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 1, we argue that an initial information advantage alone is not enough to generate the home bias. In order to make this point, we examine a model where the increasing returns to learning mechanism is shut down. An investor minimizes his portfolio variance by choosing what to learn, but takes his portfolio as given. Investors in this setting undo their natural information advantage by learning as much as possible about risk factors that they do not know much about already. With sufficient capacity, all initial information advantage, and therefore all home bias, is erased (section 1.1). To generate a large home bias, foreign information would have to be more expensive or more difficult to acquire. We quantify this cost by requiring investors to use more capacity to process foreign information (section 1.3). To generate the amount of home bias in the data, processing foreign information would have to be more than seven times harder than processing domestic information. Based on actual costs of translation (relative to the costs of financial analysis), this cost ratio strikes us as implausibly high.

Section 2 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 (long or short) 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. The force behind this is increasing returns to information: the more shares a piece of information can be applied to, the more benefit it provides.

In general equilibrium, asset prices reinforce the incentive to specialize. The assets that investors profit most from are assets for which they are more informed than the average investor. It is for these assets that they are compensated for more risk than they bear. Maximizing this excess compensation, gives investors an additional reason to specialize, reinforcing their initial information asymmetry. As information becomes more asymmetric, home bias grows.

Calvo and Mendoza (2000) argue that more scope for diversification decreases the *incentive to learn*. In contrast, our paper shows that when learning is possible, and in particular, when agents can choose what to learn about, that the *incentive to diversify* declines.

A numerical example (section 2.5) shows that learning can 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 57%, 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 (Ahearne, Griever and Warnock, 2004).

A wide 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, and Ivkovic, Sialm and Weisbenner, 2004). Section 3.1 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 3.2 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. Finally, 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 3.3 shows why geographical proximity is such a good predictor for information and equity flows: Having a home advantage in risks that both countries share operates like having a neighboring country advantage as well. This initial advantage makes learning about the neighbor more profitable, and makes trading with a neighbor more like trading with a compatriot.

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 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 agent 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."

1.1 A Model without Increasing Returns to Information

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. Prices are set such that the market clears. In period 3, he receives the asset payoffs and consumes.

The vector of unknown asset payoffs f is what the investor will devote capacity to learning about. The investors choose signals to maximize their period-2 expected exponential utility of period-3 profits.

$$U = -E_2\{exp(-\rho q'(f - pr))\} \tag{1}$$

where ρ is risk aversion, r is the risk-free return and p and q are Nx1 vectors of the asset prices and the number of shares of each asset in the investor's portfolio. Rather than regarding the portfolio as an endogenous choice variable, the investor takes q as given when choosing what to learn. It is this assumption that shuts down the increasing returns to scale. We will allow the agent to consider the effect of his learning on his portfolio in section 2.

Each investor is endowed with prior beliefs about the vector of asset payoffs $f \sim N(\mu, \Sigma)$. Home and foreign investors differ only in the accuracy of their prior beliefs. 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.

At time 1, investors choose how to allocate information capacity. When asset payoffs co-vary, learning about one payoff is informative about others. Therefore, it is useful to state the problem as learning about factors, rather than assets. We allow investors to learn about risk factors that are linear combinations of asset payoffs, proportional to the principal components of Σ. The principal components result from decomposing the prior belief variance-covariance matrix Σ, into a diagonal eigenvalue matrix Λ, and an eigenvector matrix revealing the relative loadings of each asset on each principal component: $\Sigma = \Gamma\Lambda\Gamma'$. For example, if the first principal component represented U.S. business cycle risk, then the first column of Γ would tell us about how much each asset's payoff co-varied with the U.S. business cycle, and $\Lambda_{1,1}$ would tell us about how risky, or how uncertain, U.S. business cycles are. Learning about principal component risks means that investors will have posterior beliefs with the same principal components as their prior beliefs ($\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$), but with lower weights $\hat{\Lambda}_i$ on some risks they chose to learn about, reflecting the decrease in variance of those risks, conditional on the new information.

There are 2 constraints governing how the investor can choose his signals. The first constraint governs the total capacity the investor is allowed to use to transmit information. Our measure of information *capacity* is the standard measure in information theory: the *reduction in entropy.* The entropy of a variable $\chi \sim p(\chi)$ is $-E[\log(p(\chi))]$. For an n-dimensional multivariate normal, with variance-covariance matrix V, entropy is $\frac{1}{2}\log\left((2\pi e)^n|V|\right)$. Like variance, entropy is a measure of uncertainty about a variable. It is a stock; capacity is its flow.[1] The more information the signal contains, the more the posterior variance falls below the prior variance, and the more information capacity is required to transmit the signal. The capacity constraint states that the total information capacity required to achieve posterior belief $\widehat{\Sigma}$ must not exceed a given scalar K:

$$\frac{1}{2}\left[\log(|\Sigma|) - \log(|\widehat{\Sigma}|)\right] \leq K \tag{2}$$

The second constraint is that 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 choosing to erase prior information, does not make sense in this context, we rule this out by requiring the signal variance-covariance matrix to be positive

[1] In statistics, the Kullback-Liebler distance is used as a measure of how difficult it is to distinguish a distribution from an uninformative prior. Another way to interpret capacity is that it constrains the increase in this distance, after observing the chosen information.

semi-definite. That restriction implies that

$$\Sigma - \widehat{\Sigma} \quad \text{positive semi-definite.} \tag{3}$$

Conditional on seeing signal $\eta \sim N(f, \Sigma_\eta)$, the investor's period-2 posterior belief about the asset payoff is $f \sim N(\hat{\mu}, \widehat{\Sigma})$.[2] Using standard Bayesian updating formulas, the posterior mean and variance can be expressed as

$$\hat{\mu} \equiv E[f|\mu, \eta] = \left(\Sigma^{-1} + \Sigma_\eta^{-1}\right)^{-1} \left(\Sigma^{-1}\mu + \Sigma_\eta^{-1}\eta\right) \tag{4}$$

$$\widehat{\Sigma} \equiv V[f|\mu, \eta] = \left(\Sigma^{-1} + \Sigma_\eta^{-1}\right)^{-1}. \tag{5}$$

Without loss of generality, we bypass the choice of signals and model the choice over the posterior beliefs directly. 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 of each risk factor to face. We rewrite period-2 expected utility to eliminate the 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, and substituting $\Gamma\hat{\Lambda}\Gamma$ for $\widehat{\Sigma}$, we can restate the optimal learning problem as:

$$\max_{\hat{\Lambda}} - \exp\left(-\rho q'(\hat{\mu} - rp) + \frac{\rho^2}{2} q'\Gamma\hat{\Lambda}\Gamma' q\right) \quad \text{s.t. (2) and (3).} \tag{6}$$

1.2 Results: Learning Undoes Information Advantages

Let $\tilde{q}_i$ be the ith entry of the vector Γq. This represents the amount of risk factor i that an agent holds.

Proposition 1 *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. This type of solution is called a "water-filling" result in the information theory literature because the way in which capacity is allocated resembles the way water fills up bins of different depth. In figure 1, each bin corresponds to a principal component, or risk factor

[2] Choosing normally distributed signals is optimal because normal distributions maximize the entropy over all distributions with a given variance (see Cover and Thomas (1991), chapter 10).

(an eigenvector of Σ), that the investor can learn about, and the depth of the bin is determined by the amount of that risk present in his portfolio $\tilde{q}_i^2$.

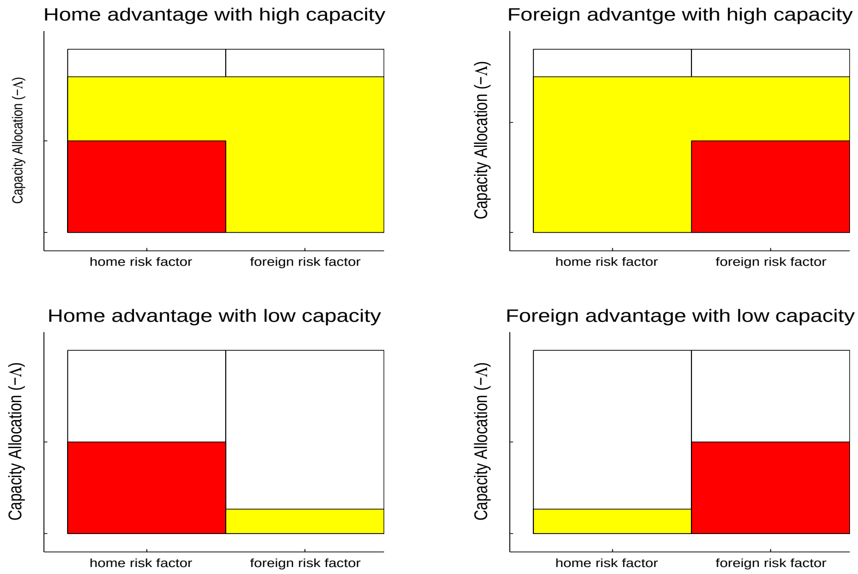

Figure 1: Allocation of information capacity for a low and high-capacity representative investor. The lightly shaded area represents the amount of capacity allocated to the factor. The dark area represents the size of the information advantage. The unfilled part of each bin represents the posterior variance 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.

Because the investor has a fixed target $(\frac{1}{\tilde{q}_i^2}M)$ for how much he wants to learn about each risk factor, whenever he is given an initial information advantage in one risk factor, he compensates for it by choosing to learn more about other risk factors. 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. 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 information advantage (the brick). Giving an investor a home information advantage is like placing a brick in the left side of the box. Giving him a foreign advantage is like placing the brick on the right side. When capacity is high, a sufficiently small brick placed on either side will raise the water level on both sides equally. The home investors shifts capacity to the foreign risk and the foreign investor learns more about the home risk. 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, where the initial advantage is placed does not matter. With sufficient capacity, initial information advantages cannot contribute to a home bias. A home bias could still arise if an agent anticipated holding a lot of the home assets (high $\tilde{q}$ for the home assets). But the corollary states that this would not be due to the initial information advantage. The next section provides a rationale for why home investors expect to hold a lot of the home asset (section 2).

1.3 Mechanisms to Preserve Information Advantages

Without increasing returns, there are two ways that initial information advantages can persist. The first possibility is that investors capacities are less than K^*. When capacity is low or the advantage is large (top panels of figure 1), the brick raises the level of the particular bin it is thrown in to – it reduces the posterior variance of that risk. This gives us a candidate explanation for information immobility: investors have large information endowments and learning is difficult. However, if this explanation were true, then individuals would never choose to learn about local assets; they would devote what little information capacity they have entirely to learning about foreign assets. This implication is inconsistent with the multi-billion-dollar industry that analyzes U.S. stocks, produces research reports on the U.S. economy, manages portfolios heavy in U.S. assets, and then sells their products, in large part, to Americans. Furthermore, in a setting without increasing returns to information, Pastor (2000) shows that even an investor who makes no learning choices, but passively observes all return realizations ($K = 0$), must have implausibly precise prior beliefs to justify the observed home bias.

A second candidate explanation is that investors have a harder time processing information about foreign assets. If analyzing foreign information consumes more capacity (perhaps because of language barriers), investors might rationally choose to learn more about home information and reinforce information asymmetry. Two crucial questions arise: How much more difficult must it be to learn foreign information in order for this to explain the degree of home bias we observe? Is the implied magnitude of this foreign information friction realistic?

To explore the first question, we replace (2) with a capacity constraint that requires more capacity to process foreign than home information. Next, we look at the optimal learning choice and the resulting optimal portfolio. Then, we ask how difficult foreign information must be to learn in order to explain the fact that U.S investors hold only 12% foreign assets. The level of difficulty can be compared to costs of obtaining processed and translated home and foreign information.

We investigate a simple setting with one home and one foreign asset, with variances σ_h^2 and σ_f^2 and zero covariance.[3] Then, the capacity constraint with asymmetric information processing costs is:

$$\frac{1}{2}\left[\log(\sigma_h) - \log(\hat{\sigma}_h)\right] + \frac{\psi}{2}\left[\log(\sigma_f) - \log(\hat{\sigma}_f)\right] \leq K \tag{7}$$

where h and f subscripts denotes prior and posterior variances for home risks and foreign risks. ψ is then the ratio of capacity that must be devoted to foreign and to home risk to obtain the same amount of information.

Taking first order conditions with respect to σ_h^2 and σ_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 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 $\frac{q_h^\star}{q_f^\star} = \frac{(\hat{\sigma}_h^2)^{-1}}{(\hat{\sigma}_f^2)^{-1}} = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_h^2}$. The average U.S. investor holds 88% home and 12% foreign assets; that is 7.3 times more home assets than foreign assets. Such a portfolio implies that ψ must be at least 7.3.[4]

If home investors start with an information advantage in home assets, might this reduce the required cost of foreign information? The answer to this question is no. The initial information

[3]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, the fact that nothing is learned about foreign assets from a signal about the home asset will cause the model to overestimate the home bias. This will bias ψ downward.

[4]It is possible that the investor starts out believing that $q_h > q_f$, reducing the required processing cost. But then the home bias would not be wholly due to the processing cost, but would be partly because of expected portfolio holdings. This expectation of q_h is central to the theory explored in the next section.

advantage only affects whether or not it is feasible to achieve a posterior variance ratio of 7.3. Lower initial home variance requires less capacity to reach this target ratio. As long as the investor does not start with beliefs that are more than 7.3 times more precise for home assets, his learning choices can only get him to the target ratio if $\psi \geq 7.3$.

The model's predicted relative *shadow* price of foreign information ($\psi = 7.3$) seems out of line with the relative *market* price of foreign information. For individual investors, foreign financial newspapers, even ones translated in English, are readily available for high market-value countries like Germany, France, Spain, Italy and of course, the UK. For investment firms, average salaries for translators are typically 25% less than for financial analysts.[5] 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 six times the analyst's to generate the required cost ratio. Looking as cost of translation measured in terms of the volume of material to be translated leads us to the same conclusion. A 5000-page report costs approximately \$900 to translate.[6] If differences in information costs were the cause of the home bias, translation would have to be six times as expensive as analysis, meaning that the price of a 5000-page research report by a home analyst must cost no more than \$150. It is possible that agency problems, legal hurdles, and differences in reporting standards prevent some international trade in information. But the magnitude of the hurdles must be implausibly large.

The model discussed in this section shuts off the increasing returns to information mechanism proposed in the introduction, by holding investors' portfolios fixed when they choose what to learn. Without this mechanism, generating beliefs that differ greatly across countries is an uphill battle. When initial information advantages are not too large, or learning capacity is high, the effects of initial advantages disappear. If difficulty processing foreign information were the cause of home bias, the cost ratio for foreign to home information needs to be very large. If home investors can easily learn about foreign information, but hold under-diversified portfolios, is it possible that they optimally choose not to learn foreign information? Maybe learning more about what they already have an advantage in is optimal.

[5] 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.

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

2 A Rational Expectations Model of Specialized Learning

This section analyzes a decentralized model where small information advantages 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 agent takes into account that every portfolio in the market depends on what each investor learns. Ultimately, we conclude that the assumption of information immobility is a defensible one, but not for the reasons originally thought. 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.

We model two countries, home and foreign. Each has an equal-sized continuum of investors, whose preferences are identical and given by (1). 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 $\mu^\star \sim N(f, \Sigma^\star)$. In period 1, each investor chooses a distribution from which they will draw signals about asset payoffs, subject to their capacity constraint. In period 2, investors observe signals from the distributions they have chosen. We assume that the draws from these distributions are *independent* across agents. The independent noise assumption can also be thought of as an independent error that each investor adds when he interprets his information. We believe that this is the relevant physical constraint that humans are facing when trying to process financial information (see Sims 2003). In period 2, investors also observe prices. When agents' portfolios were fixed (section 1), what investors learned could not affect the market price. Now that asset demand responds to observed information, the market price will be a noisy signal of this aggregated information. Using their prior beliefs, their chosen signals, and information contained in prices, the investors form asset portfolios. Finally, in period 3, asset payoffs are realized and investors consume their profits.

2.1 The Period-2 Portfolio Problem

We solve the model using backwards induction, starting with the optimal portfolio decision, taking information choices as given. In section 2.2, we look at what information investors will choose to learn. In period 2, investors have three pieces of information that they must aggregate to form their expectation of the assets' payoffs: their prior beliefs, their signals (draws from distributions chosen in period 1), and the equilibrium asset prices.

Given posterior mean $\hat{\mu}_j$ and variance $\widehat{\Sigma}_j$ of asset payoffs, the portfolio for investor j, from either country, is

$$q_j = \frac{1}{\rho}\widehat{\Sigma}_j^{-1}(\hat{\mu}_j - pr). \tag{8}$$

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 (nx1) 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

$$\int_0^1 \widehat{\Sigma}_j^{-1}(\hat{\mu}_j - pr)dj = \bar{x} + x. \tag{9}$$

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. The pricing formula is slightly different from Admati (1985) and Van Nieuwerburgh and Veldkamp (2004) because prior beliefs are not common across agents and therefore do not introduce a common shock to price signals.

Belief updating If prices take this form, then the mean and variance of the asset payoff, conditional on prices are $E[f|p] = (rp - A)$ and $V[f|p] = \sigma_x^2 C'C \equiv \Sigma_p$. The posterior belief about the asset payoff f, conditional on a prior belief μ_j, signal $\eta_j \sim N(f, \Sigma_{\eta j})$, and prices, can be expressed using standard Bayesian updating formulas:

$$\hat{\mu}_j \equiv E[f|\mu_j, \eta_j, p] = \left(\Sigma_j^{-1} + \Sigma_\eta^{-1} + \Sigma_p^{-1}\right)^{-1}\left(\Sigma_j^{-1}\mu_j + \Sigma_{\eta j}^{-1}\eta_j + \Sigma_p^{-1}B^{-1}(rp - A)\right) \tag{10}$$

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

$$\hat{\Sigma}_j \equiv V[f|\mu_j, \eta_j, p] = \left(\Sigma_j^{-1} + \Sigma_{\eta j}^{-1} + \Sigma_p^{-1}\right)^{-1}. \tag{11}$$

These are the conditional mean and variance that agents use to form their portfolios in (8).

In order to determine the amount of capacity used in learning, we will need to know what variance of payoffs an investor with zero capacity would face. Since that investor can learn from prices and prior beliefs, he faces a posterior variance that we will call $\tilde{\Sigma} \equiv V[f|\mu_j, p] = \left(\Sigma_j^{-1} + \Sigma_p^{-1}\right)^{-1}$.

2.2 The Optimal Learning Problem

To make aggregation tractable, we assume that Σ and $\Sigma^\star$ have the same eigenvectors, but different eigenvalues. As in the previous section, we assume that investors can observe signals proportional to these eigenvectors (principal components), subject to their capacity constraint. 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. Our baseline analysis will 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. Section 3 will relax the symmetry assumption. Then prior beliefs are

$$\Sigma = \Gamma\Lambda\Gamma' \tag{12}$$

$$\Sigma^\star = \Gamma\Lambda^\star\Gamma' \tag{13}$$

where $\Lambda^\star = \Lambda \begin{bmatrix} (1+\alpha)I_{N/2} & 0 \\ 0 & \frac{1}{1+\alpha}I_{N/2} \end{bmatrix}$, and $\alpha \geq 0$ measures the strength of the information advantage.

We state the period-1 information choice problem as choosing a signal distribution to maximize the period-1 expectation of period-2 expected utility. The prices realized in period 2 are now random variables in period 1. The crucial difference between this model and the one in section 1 is that here, each investor chooses a portfolio that depends on his observed signal. Substituting in the expression for the equilibrium asset demand from the period 3 problem (8), into the left-hand side of (14) produces an expected utility function that is a moment generating function of a quadratic form of the normal variable $(\hat{\mu}_j - pr)$.

$$-E_1\left[E_2\left[\exp\left(-\rho q_j(f-pr)\right)\right]\right] = -E_1\left[exp\left[-\frac{1}{2}(\hat{\mu}_j - pr)'\widehat{\Sigma}_j^{-1}(\hat{\mu}_j - pr)\right]\right] \tag{14}$$

As before, each investor can choose to learn about risk factors Γ that are eigenvectors of the variance-covariance matrix of asset payoffs that they would face if they had zero capacity to learn.

That zero-capacity variance is no longer Σ, now it is $\tilde{\Sigma}$, which contains information from prior beliefs and prices. At time 1, $(\hat{\mu}_j - pr)$ is a normal variable, with mean $(-A)$ and variance $\Sigma_p - \widehat{\Sigma}_j$.[7] Using the moment generating function formula (see Mathai and Provost, 1992), we can express the period-1 optimization problem of an investor:

$$\max_{\widehat{\Sigma}_j} \; \log\left(|\Sigma_p - \widehat{\Sigma}_j|\right) - \log\left(|\widehat{\Sigma}_j|\right) + A'(2\widehat{\Sigma}_j^{-1} - (\Sigma_p - \widehat{\Sigma}_j)^{-1})A. \tag{15}$$

The choice of the covariance matrix of the posterior belief $\widehat{\Sigma}$ is subject to two constraints. The first constraint is that the information the investor sees cannot reduce entropy by more than a fixed capacity K:

$$\frac{1}{2}\left[\log(|\tilde{\Sigma}_j|) - \log(|\widehat{\Sigma}_j|)\right] \leq K \tag{16}$$

This information processing constraint is slightly different from the one in (2): It is still a constraint on the distance between $\widehat{\Sigma}$ and a reference variance, but now the reference variance is $\tilde{\Sigma}$, instead of Σ. $\tilde{\Sigma}_j = V[f|\mu_j, p]$ is what the conditional variance of asset payoffs would be if the agent only knew what was in his prior beliefs (including his information advantage), and what he observed from the price level. In both models, the reference variance is the conditional variance of asset payoffs that an investor with zero capacity faces.

The second constraint is the equivalent of (3). It prevents the investor from acquiring negative information.

$$\tilde{\Sigma}_j - \widehat{\Sigma}_j \quad \text{positive semi-definite} \tag{17}$$

2.3 Results: Learning with Increasing Returns

Consider the eigenvalue decomposition of the 'pre-signal' variance matrix $\tilde{\Sigma}_j = \Gamma\tilde{\Lambda}_j\Gamma'$. Note that this variance matrix has the same eigenvectors Γ as the prior beliefs.[8] Let $\hat{\Lambda}^a = (\int_j \hat{\Lambda}_j^{-1})^{-1}$ be the eigenvalue matrix of a hypothetical investor whose posterior belief precision is the average of all investors' precisions. The ratio of $\hat{\Lambda}^a$ and agent j's $\tilde{\Lambda}$ tells us, for each risk factor, how much capacity would be required for j to become as well-informed as the average investor .

Proposition 3 ***Optimal Information Acquisition*** *In general equilibrium with a continuum of*

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

[8] To see why this is true, note that $\tilde{\Sigma}$ is the harmonic mean of prior belief variance and Σ_p. From appendix A.2 we know that $\Sigma_p = \frac{1}{\rho^2\sigma_x^2}\Gamma\Lambda'_{\eta a}\Lambda_{\eta a}\Gamma$, where the diagonal matrix $\Lambda'_{\eta a}\Lambda_{\eta a}$ represents the squared eigenvalues of $\int_j \Sigma_{\eta j}^{-1} dj$. Then, $\tilde{\Sigma}_j^{-1} = \Gamma\Lambda_j^{-1}\Gamma + \Gamma(\Lambda'_{\eta a}\Lambda_{\eta a})^{-1}\Gamma$. Therefore, $\tilde{\Sigma}_j = \Gamma(\Lambda_j + \Lambda'_{\eta a}\Lambda_{\eta a})\Gamma$, which has eigenvectors Γ.

investors, each investor's optimal information portfolio uses all capacity to learn about one linear combination of asset payoffs. The linear combination is the eigenvector Γ_i associated with the highest value of $\tilde{\Lambda}_i^{-1}(\hat{\Lambda}_i^a)^2(\Gamma_i'\bar{x})^2$.

Proof: See appendix A.3.

This result tells us that three features make a risk factor desirable to learn about. First, an investor should learn about a risk that is abundant (high $(\Gamma_i'\bar{x})^2$). Since information has increasing returns, the investor gains more from learning about a risk that is present in many large assets. Second, the investor should learn about a risk factor with a high expected return. Returns to a given risk increase when the average investor bears more of that risk. (Appendix A.2 shows that $E[f - pr] = \rho\Gamma\hat{\Lambda}^a\Gamma\bar{x}$.) 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 $\hat{\Lambda}_i^a/\tilde{\Lambda}_i$). Since these are the assets he will expect to hold more of, these are more valuable to learn about.

Another interpretation of this learning rule is that it tells the investor to learn about the risk factor with the highest squared Sharpe ratio. The expected return on risk i is $\Gamma_i'E[f - pr]$. Its period-2 standard deviation is $\tilde{\Lambda}_i^{1/2}$ if the investor does not learn about i, and $e^{-K}\tilde{\Lambda}_i^{1/2}$ if he does.

This result also tells us about what the cross-sectional distribution of learning will be. Because investors affect each others' learning decisions by lowering the $\hat{\Lambda}_i^a$ for the risk they learn about, the aggregate allocation is a Nash Equilibrium. Consider constructing this Nash equilibrium by an iterative choice process. Each investor will begin by choosing the largest $\tilde{\Lambda}_i^{-1}(\tilde{\Lambda}_i^a)^2(\Gamma_i'\bar{x})^2$. Suppose capacity is high and there is another risk factor j whose quantity $\Gamma_j\bar{x}$ is not far below the quantity of i. Then the fall in $\hat{\Lambda}_i^a$, brought on by some investors learning about i will convince other investors to switch to learning about j. In equilibrium, 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. The number of home risk factors learned about, in the aggregate, will depend on the amount of capacity home investors have and the relative sizes of the home risk factors. Despite the fact that many risk factors are potentially being learned about in equilibrium, it remains true that each investor concentrates all his capacity on learning about one of these factors.

Home investors always get more value from learning about home risks than foreigners do. Likewise, foreigners get higher utility than home investors do from learning about foreign risks.

Some foreigner may choose to learn about home risks, if home risks are particularly valuable to learn about. 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. Therefore, investors never make up for their initial information asymmetry by each learning about the others' advantage. They either preserve the gap in knowledge, or amplify their differences. That is the content of the following corollary.

Let $\Lambda_h, \Lambda_f, \hat{\Lambda}_h$ and $\hat{\Lambda}_f$ be $N/2$-by-$N/2$ diagonal matrices that lie on the diagonal quadrants of the prior and posterior belief matrices: $\Lambda = [\Lambda_h 0; 0 \Lambda_f]$ and $\hat{\Lambda} = [\hat{\Lambda}_h 0; 0 \hat{\Lambda}_f]$. And, let the * superscript on each of these matrices denotes foreign belief counterparts. Then, for example, $\log(|\Lambda_f|)$ represents home prior belief uncertainty (entropy) due to foreign risk factors and $\log(|\hat{\Lambda}^\star_h|)$ represents foreign posterior uncertainty, due to home risk factors.

Corollary 2 ***Learning Amplifies Initial Information Advantages*** *Suppose that capacity is positive $K > 0$ and equal for all investors, then learning will amplify initial differences in prior beliefs:* $\log(|\Lambda^\star_h|) - \log(|\Lambda_h|) \leq \log(|\hat{\Lambda}^\star_h|) - \log(|\hat{\Lambda}_h|)$ *and* $\log(|\Lambda_f|) - \log(|\Lambda^\star_f|) \leq \log(|\hat{\Lambda}_f|) - \log(|\hat{\Lambda}^\star_f|)$, *for every pair of home and foreign investors.*
Furthermore, if home and foreign risk factors are symmetric: for every home factor i, there is a foreign factor j such that $\Lambda_i = \Lambda^\star_j$ and $\Gamma_i \bar{x} = \Gamma_j \bar{x}$, then home investors will learn exclusively about home risks and foreign investors will learn exclusively about foreign risks.

Proof: See appendix A.4. When both sets of risk factors are symmetric, the value of learning about home or foreign risk factors (the first $N/2$ or the last $N/2$ Γ_i's) is determined by the ratio of the average investor's $(\hat{\Lambda}^a_i)^2$ to the investor's $\tilde{\Lambda}_i$. Recall that initial information advantage was such that $\tilde{\Lambda} < \tilde{\Lambda}^\star$ for home risks, and vice-versa for foreign risks. That means that for every home investor who considers learning about a foreign risk, there is a foreign investor who has a stronger incentive to learn about that risk. Foreign investors continue to push down the average amount of that risk $\hat{\Lambda}^a_i$, until the expected return is so low that the home investor no longer wants to learn about that risk. This proposition holds for *any* initial advantage $\alpha > 0$, it can amplify the smallest information advantage.

When the size and variance of risk factors are not symmetric, some investors may learn about risks they have no advantage in. But, neither group will ever learn more than the other group about assets that the other has an advantage in. We will explore the effect of relaxing the symmetric capacity assumption in section 3.

2.4 Home Bias in Investors' Portfolios

What do these learning strategies imply for the size of the home bias? To answer this question, we need to fix a definition of the benchmark diversified portfolio. We consider two benchmarks. The first portfolio is one with no information advantage and no capacity to learn (α=0 and $K = 0$). 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 $\bar{x}$.

$$E[q_j^{no\ adv}] = \bar{x} \ \ \forall j \tag{18}$$

A second natural benchmark portfolio is one where investors have initial information advantages, but no capacity to learn ($\alpha > 0$, $K = 0$). This is the kind of information effect that Ahearne, Griever and Warnock (2004) estimate when they look at how much of the home bias could be accounted for by home investors' lack of knowledge of foreign accounting standards, without giving investors an opportunity to learn.

$$E[q_j^{no\ learn}|\mu_j, p] = \Gamma\tilde{\Lambda}_j^{-1}\hat{\Lambda}^a\Gamma'\bar{x}. \tag{19}$$

When investors have capacity to learn, their realized portfolios will depend on the exact realization of the prior belief they draw, the realization of the signal they draw, as well as on the realization of the asset supply shock working through the price level. Let Υ be an NxN identity matrix, except for the (i,i)th element, which is e^{-2K}, where i =argmax$\frac{(\hat{\Lambda}_i^a)^2}{\tilde{\Lambda}_i}\Gamma_i'\bar{x}$. Then, investor j, with an initial information advantage $\alpha < 1$ and capacity to learn $K > 0$, holds the expected portfolio

$$E[q_j|\mu_j, \eta_j, p] = \Gamma(\tilde{\Lambda}_j\Upsilon_j)^{-1}\hat{\Lambda}^a\Gamma'\bar{x} \ \ \forall j \tag{20}$$

The specialization in learning does not imply that the investor holds exclusively assets along the risk dimension he learns about. Investors still exploit gains from diversification; they hold risks they do not learn about. Learning choices twist the portfolio away from the fully diversified portfolio in (18), towards the risk factor they learn about. The more capacity, the more pronounced the twist. The next proposition formalizes the difference between the portfolio in (20) and the fully diversified portfolio in (18). Let Γ_h represent the first $N/2$ columns of the eigenvector matrix Γ. It maps assets into loadings on home risk factors.

Proposition 4 ***Information Mobility Increases Home Bias*** *When home and foreign risk factors are symmetric, a home investor with an information advantage $\alpha > 0$ holds more of assets that load on home risk when he can learn ($K > 0$), than when he cannot ($K = 0$): $\Gamma_h' E[q] \geq \Gamma_h' E[q^{no\ learn}]$.*

Proof: See appendix A.5. There are two forces causing investors to overweight home assets in their portfolios: an information advantage and a general equilibrium effect. The information advantage is the fact that home investors start out with more information about those assets, find it optimal to specialize in learning more about home assets, and therefore conditional on their information, they face less risk with home assets. (Diagonal entries $(\Lambda_i \Upsilon_i)$ are low for home assets i.) The general equilibrium effect arises because foreigners have an information advantage in foreign assets. Facing less risk in foreign assets, foreigners demand more of the foreign assets and push up their price. Not only do foreign assets appear more risky to home investors because of their initial information disadvantage, but their expected returns are low because foreign investors' learning has driven up the price. In sum, foreign assets do not provide enough reward to home investors, relative to the risk they carry.

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 5 ***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 a) takes a larger position (long or short) in the assets that he learns about, and b) increases the correlation of his asset demand q with asset payoffs $(f - pr)$.[9] He holds

[9] $\text{Corr}(q, f - pr) = \text{corr}(\hat{\mu} - pr, f - pr)$. Since, $\hat{\mu}$ is an unbiased expectation of f, it is equal to $f + \epsilon$, where ϵ is

a long position in the asset when it is likely to pay high returns, and shorts the asset when it is likely to pay low returns.

2.5 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 2. There are two countries, home and foreign, with a large number of agents (1000) in each. The risk aversion parameter of all agents is $\rho = 2$. There are N assets in the economy, $N_h = 5$ home assets and $N_f = 5$ foreign assets.

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 (risk from noise traders) affects how informative prices are. It 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 average agent's prior belief is equal to the true asset payoffs. The prior variance is constructed in two steps. First, the standard deviation of payoffs is chosen between 5-10%, such that all assets have the same expected payoff to standard deviation ratio. Second, the variance of an investor's home assets is reduced by a factor $\alpha = 0.1$. For each country, the home agents' prior beliefs about home assets have a variance that is ten percent lower than foreign agents' variance. The reverse is true for foreign assets. We vary learning capacity K to explore its effect.

Following convention, we define the home bias as

$$\text{home bias} = 1 - \frac{1 - share\ of\ home\ asset\ in\ home\ portfolio}{share\ of\ foreign\ assets\ in\ world\ portfolio} \tag{21}$$

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

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.

advantage ($\alpha = 0.1$), but no learning capacity (K=0). Still, home investors can partially learn the foreign investors' information advantage through prices. The ten percent initial information advantage leads to a 4.6 percent home bias in asset holdings. Next, we switch on the learning mechanism by giving each agent 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 eliminate 20 percent of the risk in one asset (K=.22, $1 - e^{-K} = .20$), the home bias doubles from 4.6 to 9.3 percent. When there is enough capacity to eliminate 57 percent of the risk in one asset (K=.85, $1 - e^{-K} = .57$), the home bias is 31 percent, almost eight times as large as without learning effect. A thirty percent home bias is large, because in an economy with uncorrelated assets, the gains from diversification are very strong.

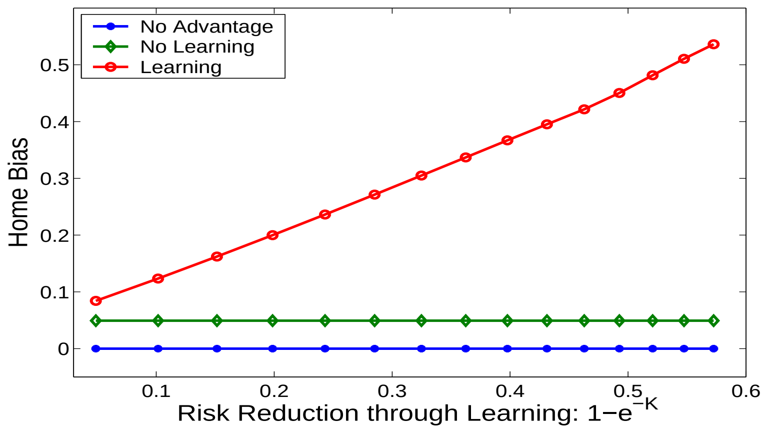

Figure 2: Home Bias Increases With Capacity.
Numerical example with two countries. Assets within a country have correlated payoffs (cov= $.03^2$). Home bias is defined in (21). The 'no advantage' line (stars) gives the home bias in an economy with no initial informational advantage and no capacity to 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 0.75. 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 have the same correlation structure as home assets, each agent learns about one risk factor, that all his domestic assets load on. Introducing correlation almost doubles the home bias: Home bias is 20% when $1-e^{-K} = .20$ and increases to 54% for $1-e^{-K} = .57$. (See line with circles in figure 2.) In contrast, the no learning benchmark is virtually unaffected (4.7%, line with diamonds). There are two reasons for this increase. First, 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. Second, every home asset's price provides information about the asset payoff of each other home asset. Learning from prices is easier in such a setting. The higher the precision of beliefs after learning from prices ($\tilde{\Lambda}^{-1}$), the more of an increase in precision through learning a given capacity allows ($\hat{\Lambda}_i^{-1} = \tilde{\Lambda}_i^{-1} e^{2K}$). This second channel acts like an increase in capacity.

3 Extensions

3.1 Local Investing: Heterogeneous Home Information

In the model of section 2, 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 agent j from region k draws an independent prior belief from the distribution $\mu_j \sim N(f, \Sigma_k)$, where $\Sigma_k = \Gamma \Lambda_k \Gamma$, and Λ_k has a kth diagonal entry that was lower than the kth diagonal in the beliefs of any other region. With this setup, all of the results of section 2 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 5). 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.

Numerical Example To quantify the local bias, we use the same numerical example as in section 2.5. 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 2.5 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. I.e., the aggregate information advantages at home and abroad are unchanged. We measure local bias as:

$$\text{local bias} = 1 - \frac{1 - share\ of\ local\ asset\ in\ local\ portfolio}{share\ of\ non-local\ assets\ in\ world\ portfolio} \tag{22}$$

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 4.4%. With learning capacity, K=.85, the average local bias is 19%. This average understates the local bias in many regions; the highest is 50%. A local bias of 19% (50%) implies that the local investor holds 27% (55%) of their portfolio in the local asset. This is 2.7 (5.5) times larger than what full diversification would predict.

Local bias results illustrate the effect of asymmetry in factor size and variance. Recall that the different home assets have different variances (Λ_i) and different quantities of risk ($\Gamma_i \bar{x}$). As a result, investors in some regions choose to specialize in a larger or higher-return risk factor, in another region. When local investors choose not to learn about local risk, nobody will learn about it. (See corollary 2.) Hence, the region's investors will retain their initial information advantage about the local asset. Since they are more informed than the average investor, they hold more than the diversified share of the local risk factor in their portfolio. Local bias in such regions will be positive, but small. In other regions, investors will reinforce their local advantage through learning. These are regions where local bias is large.

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 11% for low capacity ($K = 0.05$) and 69% for high capacity ($K = 0.85$). We can regenerate the 76% home bias observed in the data by giving investors capacity of $K = 0.8$ and a 70% local advantage.

Giving investors a local advantages as opposed to a home advantage in all home assets (section 2.5) generates 15% more home bias (for $K = .85$), even though the aggregate information advantage is the same. Why is this? An investor's portfolio share in an asset depends on the difference between the risk he bears and the risk he is compensated for. The latter is a function of the risk the average investor bears ($\hat{\Lambda}^a$). When fewer investors share an advantage in the same risk, they are more advantaged relative to the average investor, and earn higher returns. Because the returns to specialization are higher, investors diversify their portfolios less.

3.2 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: $\hat{\Lambda}^a_i < \hat{\Lambda}^{a\star}_j$, for equally-sized home and foreign risks i and j. 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, reducing 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.[10] 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 5, 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 j.

What are the Taiwanese risk factors j that Americans will learn about? Assuming that the average uncertainty ($\hat{\Lambda}^a$), and American uncertainty ($\tilde{\Lambda}$) 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.[11] 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.

3.3 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 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

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

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

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 are consistent with an information explanation, if learning was concentrated in nearby countries, or countries with highly correlated returns.

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 the Portes-Rey-Oh gravity model 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. (See 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.

4 Conclusions

This paper has 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. 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, modelling 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) have shown that human capital and equity returns are highly correlated across countries. Although this correlation worsens the home bias puzzle in a standard model, it bolsters our explanation. Labor income is a source of initial information advantage; it also provides an incentive to learn more about home assets, which are informative about future labor income realizations.

An important assumption in our model was that every agent must process their own information. A relevant question is whether one portfolio manager could process information for many investors at once. While efficiency would suggest centralization, the fact that investors want to learn what others don't know suggests an equilibrium that is neither fully centralized nor fully decentralized. 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. The 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 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] Coval, Joshua. "International Capital Flows when Investors Have Local Information." Working paper, 2000.

[9] 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.

[10] 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.

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

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

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

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

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

[16] 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.

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

[18] Mathai, Arak and Provost, Serge. *Quadratic Forms of Random Variables: Theory and Applications.* Marcel Dekker, New York, 1992.

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

[20] 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.

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

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

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

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

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

[26] 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.

[27] Van Nieuwerburgh, Stijn and Veldkamp, Laura. "Information Acquisition and Portfolio Under-Diversification." Working paper, 2004.

A Proof of Proposition 1

The optimization problem is

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

s.t. $\hat{\Lambda}_i \leq \Lambda_i$ and $\prod_i \hat{\Lambda}_i \geq \prod_i \Lambda_i e^{-2K}$. The first-order condition for this problem is

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

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. The result follows from the fact that $\phi_i = 0$ when $\hat{\Lambda}_i > \Lambda_i$. $\square$

A.1 Proof of Corollary 1

Proposition 1 tells us that posterior beliefs $\hat{\Lambda}_i$ are unaffected by an ϵ reduction in the prior belief $\Lambda_i - \epsilon$ whenever $\Lambda_i - \epsilon > \frac{1}{\tilde{x}_i^2} M$. For all posterior beliefs to be invariant to an ϵ information advantage, capacity must be large enough so that $(\Lambda_i - \epsilon)\tilde{x}_i^2 > M$ for $i = argmin_j (\Lambda_j - \epsilon)\tilde{x}_j^2$. Let $K^\star$ be the level of capacity such that $\min_i \left((\Lambda_i - \epsilon)\tilde{x}_i^2\right) = M$. Since the capacity used learning about a factor j is $\log(\Lambda_j) - \log(\hat{\Lambda}_j) = \log(\Lambda_j) - \log(\frac{1}{\tilde{x}_j^2} M) = \log(\Lambda_j \tilde{x}_j^2) - \log\left(\min_i \left((\Lambda_i - \epsilon)\tilde{x}_i^2\right)\right)$. Therefore, the total capacity required is

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

A.2 Proof of Proposition 2

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

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

$$B = I$$

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

$(\Sigma_\eta^a)^{-1}$ is the average precision of agents' information advantage, plus the average precision of the information they choose to learn: $(\Sigma_\eta^a)^{-1} = \frac{1}{2}\Sigma^{-1} + \frac{1}{2}(\Sigma^\star)^{-1} + \int_j \Sigma_{\eta j}^{-1} dj$, where $\Sigma_{\eta j}$ is the variance-covariance matrix of the signals that agent j observes.

Note that $\left(\frac{1}{\rho^2\sigma_x^2}(\Sigma_\eta^a\Sigma_\eta^{a\prime})^{-1} + (\Sigma_\eta^a)^{-1}\right) = \Gamma\hat{\Lambda}_a^{-1}\Gamma'$ is the average of all investors' posterior belief precisions, and that $\frac{1}{\rho^2\sigma_x^2}(\Sigma_\eta^a\Sigma_\eta^{a\prime})^{-1} = \Sigma_p^{-1}$.$\square$

A.3 Proof of Proposition 3

Let $\tilde{\theta} = \Gamma^{-1}((I-B)\mu - A)$, a linear transformation of period-1 expected excess returns.

Van Nieuwerburgh and Veldkamp (2004), proposition 10 shows that each investor's optimal learning strategy is to use all capacity to learn about the risk factor Γ_i with the highest $\tilde{\theta}_i^2\tilde{\Lambda}_i^{-1}$. In this setting, where $B = I$, $\tilde{\theta} = -\Gamma^{-1}A = \Gamma^{-1}\rho\Gamma\hat{\Lambda}_a\Gamma'\bar{x} = \rho\hat{\Lambda}^a\Gamma'\bar{x}$. (See proposition 2.) Therefore, the optimal risk factor to learn about is the one with the highest $\rho^2\tilde{\Lambda}_i^{-1}(\hat{\Lambda}_i^a\Gamma_i'\bar{x})^2$. Since ρ is a positive scalar, this is also the risk factor with the highest $\tilde{\Lambda}_i^{-1}(\hat{\Lambda}_i^a\Gamma_i'\bar{x})^2$.

A.4 Proof of Corollary 2

The value of learning about a home risk factor i is always greater for a home agent:

$$\frac{(\hat{\Lambda}_i^a)^2}{\tilde{\Lambda}_i}(\Gamma_i'\bar{x})^2 > \frac{(\hat{\Lambda}_i^a)^2}{\tilde{\Lambda}_i^\star}(\Gamma_i'\bar{x})^2. \tag{23}$$

Likewise, the value of learning about a foreign risk factor j is always greater for a foreign agent:

$$\frac{(\hat{\Lambda}_j^a)^2}{\tilde{\Lambda}_j^\star}(\Gamma_j'\bar{x})^2 > \frac{(\hat{\Lambda}_j^a)^2}{\tilde{\Lambda}_j}(\Gamma_j'\bar{x})^2. \tag{24}$$

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 value of $\frac{(\hat{\Lambda}_i^a)^2}{\tilde{\Lambda}_i}(\Gamma_i'\bar{x})^2$. This other risk factor must be a home risk factor, otherwise the foreign investor would strictly prefer to learn about it. If every home investor learns about a home risk factor, then $|\hat{\Lambda}_h| = e^{-2K}|\Lambda_h|$ and $|\hat{\Lambda}_f| = |\Lambda_f|$. By the foreign information capacity constriant, $|\hat{\Lambda}_h^\star| \geq e^{-2K}|\Lambda_h^\star|$ and $|\hat{\Lambda}_f^\star| \leq |\Lambda_f^\star|$. The same argument, in the case where one or more home investors learn about foreign risks implies that $|\hat{\Lambda}_f^\star| = e^{-2K}|\Lambda_f^\star|$, $|\hat{\Lambda}_h^\star| = |\Lambda_h^\star|$, $|\hat{\Lambda}_f| \geq e^{-2K}|\Lambda_f|$ and $|\hat{\Lambda}_h| \leq |\Lambda_h|$. Taking logs and differences of these inequalities yields the result.

Symmetric risk factors Suppose all foreign investors besides this 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. We assumed that $\Lambda_i = \Lambda_j$ for the equal-sized risk factors. Since an equal fraction of agents is learning about i and j, the average signal precision for each risk, and the precision of the price signal will be equal $(\hat{\Lambda}^a_i)^2/\tilde{\Lambda}_i = (\hat{\Lambda}^a_j)^2/\tilde{\Lambda}^\star_j$. Therefore, home investors get as much value from learning about i as foreign investors get from learning about j:

$$\frac{(\hat{\Lambda}^a_i)^2}{\tilde{\Lambda}_i}(\Gamma_i'\bar{x})^2 = \frac{(\hat{\Lambda}^a_j)^2}{\tilde{\Lambda}^\star_j}(\Gamma_j'\bar{x})^2. \tag{25}$$

Combining (23) and (25) 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 $(\hat{\Lambda}_a)^2/\tilde{\Lambda}$ 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 $\frac{\hat{\Lambda}^a_i}{\tilde{\Lambda}_i}(\Gamma_i'\bar{x})^2 < \frac{\hat{\Lambda}^a_j}{\tilde{\Lambda}^\star_j}(\Gamma_j'\bar{x})^2$. 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. $\square$

A.5 Proof of Proposition 4

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. The difference between the portfolios with and without capacity, $E[q] - E[q^{no\ learn}]$, is $\Gamma(\Upsilon^{-1} - I)\Lambda^{-1}\hat{\Lambda}^a\Gamma\bar{x}$. (See equations 19 and 20.) But, $(\Upsilon^{-1} - I)$ has zeros everywhere, except for the (i,i)th element, which is $(e^{2K} - 1) > 0$. Since $\bar{x} > 0$ by assumption, and since $\Gamma\Lambda^{-1}\hat{\Lambda}^a\Gamma$ is positive semi-definite, $E[q] - E[q^{no\ learn}] > 0$ for all assets that load on risk factor i. $\square$

A.6 Proof of Proposition 5

Using equation (20) and appendix A.2, rewrite the return on factor i as

$$E[(\Gamma_i'q)\Gamma_i'(f - pr)] = \rho\bar{x}'\widehat{\Sigma}^a\widehat{\Sigma}^{-1}\Gamma_i\Gamma_i'\widehat{\Sigma}^a\bar{x} + \rho x'\Sigma^a_\eta(I - \widehat{\Sigma}\Sigma_p^{-1})\widehat{\Sigma}^{-1}\Gamma_i\Gamma_i'\Sigma^a_\eta x.$$

Since $\widehat{\Sigma}$ and Σ_p are both symmetric, we can switch their order.

$$= \rho\bar{x}'\widehat{\Sigma}^a\Gamma\hat{\Lambda}^{-1}\Gamma'\Gamma_i\Gamma_i'\widehat{\Sigma}^a\bar{x} + \rho x'\Sigma^a_\eta(\Gamma\hat{\Lambda}^{-1}\Gamma' - \Sigma_p^{-1})\Gamma_i\Gamma_i'\Sigma^a_\eta x.$$

The only term that is changing in capacity in this expression is the (i,i) entry of $\hat{\Lambda}^{-1}$, which is multiplied by e^{2K}. Therefore, the partial derivative with respect to capacity is:

$$\partial E[(\Gamma_i'q)\Gamma_i'(f - pr)]/\partial K = 2e^{2K}\rho\left[\bar{x}'\widehat{\Sigma}^a\Gamma_i\tilde{\Lambda}^{-1}\Gamma_i'\widehat{\Sigma}^a\bar{x} + x'\Sigma^a_\eta\Gamma_i\tilde{\Lambda}^{-1}\Gamma_i'\Sigma^a_\eta x\right].$$

Since the term inside brackets is the sum of two quadratic terms, $\partial E[(\Gamma_i'q)\Gamma_i'(f - pr)]/\partial K > 0$. $\square$