# Information Immobility and the Home Bias Puzzle

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

July 27, 2006

**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 model's predictions are consistent with observed patterns of local and foreign investment.

*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, Pierre-Olivier Gourinchas, Urban Jermann, David Lesmond, Karen Lewis, Arzu Ozoguz, Hyun Shin, Chris Sims, Eric Van Wincoop, Mark Wright, and seminar participants at the Prague workshop in macro theory, Budapest SED, CEPR meetings in Gerzensee, Econometric Society, EEA, Cleveland Fed int'l macro conference, FMA, FEA, AFA, Ohio State, Illinois, Iowa, Princeton, Virginia, GWU, NYU, MIT, New York Fed, Columbia GSB, Rutgers, UCLA, UCSD, the NBER EF&G meetings and NBER summer institute in international macro for helpful comments and discussion. JEL classification: F30, F40, D82. Keywords: Home bias, asymmetric information, learning.

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. 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 immobility 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.

The model's key mechanism is the interaction between information and investment choice. To illustrate its importance, section 2 shuts down the increasing returns to specialization mechanism 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. As learning capacity increases, the share of the diversified component declines. The model generates a long position in home assets, on average, because equilibrium asset returns reflect the risk that the average investor bears. An investor who learns about home assets reduces his uncertainty about their payoffs. Asset with less uncertain payoffs are less risky investments. Since home investors bear less risk than the average investor, they earn *excess risk-adjusted home returns*. To capture the excess return, investors take positive positions in their home assets. Thus the ability to learn magnifies home bias.

A numerical example (section 3.4) shows that learning can magnify the home bias considerably. When all home investors get a small initial advantage in all home assets, the home bias is between

5 and 46%, depending on the magnitude of investors' learning capacity. When each home investor gets 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 (2005)). 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 Stulz 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. If this is true, the very highest-profit investors should be locals who can also exploit the capacity of global brokers, a prediction confirmed by Dvorak (2005).

Magnifying information advantages generates effects that resemble a familiarity bias (Huberman 2001) or a loyalty effect (Cohen 2004). But Massa and Simonov (2005) argue that familiarity effects are really information-driven: familiarity affects less-informed investors more, diminishes when the profession or location of the investor changes, and generates higher returns. Similarly, Bae, Stulz and Tan (2005) document higher-precision earnings forecasts by domestic analysts and identify a causal link with home bias.

Information advantages have been used to explain exchange rate fluctuations (Evans and Lyons (2004), Bacchetta and van Wincoop (2006)), the international consumption correlation puzzle (Coval 2000), international equity flows (Brennan and Cao 1997), a bias towards investing in local stocks (Coval and Moskowitz 2001), and the own-company stock puzzle (Boyle, Uppal and Wang 2003). The finding that information advantages are sustainable, in fact amplified, when information is mobile, sets all these theories on firmer theoretical ground.

# 1 A Model of Learning and Investing

Using tools from information theory (Sims 2003), we construct a general framework to think about learning and investment choices. This framework adds a two-country structure to Van Nieuwerburgh and Veldkamp (2005). The added investor heterogeneity allows us to identify which investors learn about and hold each asset. Without it, we could not explore cross-country portfolio differences. Section 2 examines 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, 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, 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** Investors, with absolute risk aversion parameter $\rho$, 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. We show later this is equivalent to mean-variance preferences, commonly used in portfolio theory. Appendix A.1 discusses the rationale for this utility formulation in more detail.

**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.[1] Investor's terminal wealth is then his initial wealth $W_0$, plus the profit he earns on his portfolio investments:

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

[1]The results do not depend on the existence of a risk-free asset. Suppose agents can consume $c_1$ at the investment date and $c_2$ when asset payoffs are realized. If preferences are defined over $r * c_1 + c_2$, where $r$ is the rate of time preference, the solution will be identical. The earlier consumption choice takes the place of the riskless investment choice.

**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. When 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 factors and the risk of each factor. This decomposition breaks the prior variance-covariance matrix $\Sigma$ up into a diagonal eigenvalue matrix $\Lambda$, and an eigenvector matrix $\Gamma$: $\Sigma = \Gamma\Lambda\Gamma'$. The $\Lambda_i$'s are the variances of each risk factor $i$. The $i$th column of $\Gamma$ (denoted by $\Gamma_i$) gives the loadings of each asset on the $i$th risk factor. To make aggregation tractable, we assume that home and foreign prior variances $\Sigma$ and $\Sigma^\star$ have the same eigenvectors, but different eigenvalues. In other words, home and foreign investors use their capacity to reduce different initial levels of uncertainty about the same set of risks.

Investors observe signals about risk factor payoffs ($\Gamma' f$). Nothing prevents them from learning about many risk factors. The only thing this rules out is signals with correlated information about independent risks. Learning about risk factors (principal components analysis) has long been used in financial research and among practitioners. It approximates risk categories investors might study: business cycle risk, industry, regional, and firm-specific risk. We also assume that signals are normally distributed and that each investor's signal is independent of the signals drawn by other investors.

Choosing how much to learn about each risk factor is equivalent to choosing the variance of each entry of the N-dimensional signal $\Gamma'\eta$. Since the signal is unbiased, its mean is $\Gamma' f$. Learning more about risk $i$ means lowering the variance (increasing the precision) of the $i$th entry: $\Gamma_i'\eta$. The variance of a principal component is its eigenvalue. So, reducing uncertainty about the $i$th risk factor means choosing a smaller $i$th eigenvalue of the signal variance-covariance matrix $\Sigma_\eta$. Signals about the payoffs of all assets that load on risk factor $i$ become more accurate. With Bayesian updating, each $\Sigma_\eta$ results in a unique posterior variance matrix that measures the investor's uncertainty about asset payoffs, after incorporating what he learned. Since the mapping between signal choices and posteriors is unique, information choice is the same as choosing posterior variance, without loss of generality.[2] Since sums, products and inverses of prior and signal variance matrices have eigenvectors $\Gamma$, posterior beliefs will as well. Denoting posterior beliefs with a hat, $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$, where $\Gamma$ is given and the diagonal eigenvalue matrix $\hat{\Lambda}$ is the choice variable. The decrease in risk factor $i$'s posterior variance $(\Lambda_i - \hat{\Lambda}_i)$ measures the decrease in uncertainty achieved through learning.

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

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

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

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

**Comments on the learning technology** The structure we put on the learning problem keeps it as simple as possible. But many of these assumptions can be relaxed. First, our results do not hinge on the assumption that investors learn about principal components of asset payoffs. Appendix A.2 shows that agents specialize in what they know well, for any risk factor structure. Second, our framework can incorporate heterogeneous capacity. This extension helps explain patterns of

[2] To see why learning more reduces posterior variance, think of each investor as an econometrician. He is trying to estimate the end-of-period payoffs of all assets, subject to a constraint on the total amount of data he can collect. Collecting more data on one asset reduces the standard error of his estimate for that asset's payoff. The posterior variance is that standard error, squared.

foreign investment (section 4.2). Third, endogenizing capacity does not change how that capacity is optimally allocated. When cost is any increasing function of $K$, there is a capacity endowment that produces identical outcomes.[3] Finally, a learning technology with diminishing returns and un-learnable risk will moderate, but not overturn, our results. Instead of specializing in one risk, investors may learn about a limited set of risks. But it does not change the conclusion that investors prefer to learn about what they already have an advantage in. (See Van Nieuwerburgh and Veldkamp (2005) for a proof.)

While we can relax many assumptions, it is not true that every capacity constraint preserves specialization. We use this one because it is the unique expected-returns neutral technology (see appendix A.1). It is also a common distance measure in econometrics (a log likelihood ratio) and in statistics (a Kullback-Liebler distance); it is a bound on entropy reduction, an information measure with a long history in information theory (Shannon 1948); it can be interpreted as a technology for reducing measurement error (Hansen and Sargent 2001); it is a measure of information complexity (Cover and Thomas 1991), and it has been used to describe limited information processing ability in economic settings by (Sims 2003).[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 $\mu^j$, signal $\eta^j \sim N(f, \Sigma_\eta^j)$, and prices, is formed using Bayesian updating:

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

[3]Endogenizing capacity can produce other insights. Turmuhambetova (2005) shows that modeling capacity choice can explain some features of the relationship between income and investment patterns. But that problem can be neatly decomposed from our capacity allocation problem.

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

with variance that is a harmonic mean of the prior, signal, and price variances:

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

**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 clears if investors' portfolios $q^j$ sum to the asset supply: $\int_0^1 q^j dj = \bar{x} + x$.

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

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 and no-negative-learning constraints, (3) and (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;

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:

$$\max_{q,\hat{\Lambda}} E\left[\rho q'(\hat{\mu} - rp) - \frac{\rho^2}{2} q'\Gamma\hat{\Lambda}\Gamma' q|\mu, \Sigma\right] \qquad \text{s.t. (3) and (4).} \tag{7}$$

## 2 Why Might Information Advantages Disappear?

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

### 2.1 A Model without Increasing Returns to Information

In order to shut down the investment-learning interaction, suppose the investor takes $q$ as given, when choosing what to learn. Define the amount of risk factor $i$ that an investor holds in his portfolio as $\tilde{q}_i = \Gamma_i' q$. Then the objective (7) collapses to choosing $\hat{\Lambda}_i$'s to minimize $\sum_i \tilde{q}_i^2 \hat{\Lambda}_i$, subject to the capacity constraint (3) and the no-forgetting constraint $\Lambda_i - \hat{\Lambda}_i \geq 0 \ \forall i$. Appendix B.1 derives the following optimal learning rule and its corollary.

**Proposition 1.** ***Learning Undoes Information Advantages.*** *Optimal learning about principal components* $\Gamma$ *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.*

The investor has a target posterior variance for each risk ($\frac{1}{\tilde{q}_i^2} M$), that does not depend on prior variance. With sufficient capacity, he can fully compensate for any initial information advantage, by devoting capacity to more uncertain risks. If this is the case, then no matter what information the investor starts with, he will end up with the same posterior beliefs, after learning. Any home bias that might result from the information advantage disappears when investors can learn.

**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* $k \neq i$*.*

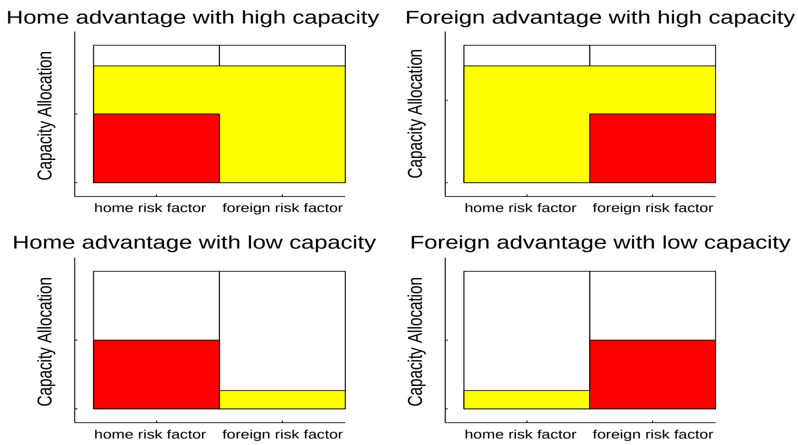


Figure 1: Allocation of information capacity for a low and high-capacity investor. The lightly shaded area represents the amount of capacity allocated to the factor. The dark area represents the size of the information advantage. The unfilled part of each bin represents the posterior variance of the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

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

The bottom panels of figure 1 illustrate low-capacity allocations. The investor would like the water level (his posterior precision) to be the same in both bins, but there is insufficient water (capacity). The no-forgetting constraint prevents him from breaking up the brick to level the water. He cannot equalize home and foreign uncertainty. The constrained optimum is to devote all capacity to the most uncertain risk.

## 2.2 Mechanisms to Preserve Information Advantages

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

A second candidate explanation is that investors have a harder time processing information about foreign assets. Consider a setting with one home and one foreign asset, with prior variances $\sigma_h^2$ and $\sigma_f^2$, posterior variances $\hat{\sigma}_h^2$ and $\hat{\sigma}_f^2$, and zero covariance. Replace (3) with a capacity constraint that requires $\psi$ times more capacity to process foreign than home information: $\frac{1}{2}[\log(\sigma_h) - \log(\hat{\sigma}_h)] + \frac{\psi}{2}[\log(\sigma_f) - \log(\hat{\sigma}_f)] \leq K$. The optimal learning choice is described by the $\hat{\sigma}_h^2$ and $\hat{\sigma}_f^2$ first order conditions. Capacity permitting, an investor sets the ratio of posterior variances to $\hat{\sigma}_f^2/\hat{\sigma}_h^2 = \psi q_h^2/q_f^2$. This implies that an investor, who initially expects to hold a balanced portfolio ($q_h = q_f$) but ends up holding 7.3 times more home assets (as in the data), must have $\psi = 7.3$.[5] Adding an initial home advantage does not alter this required processing cost, unless the advantage alone can account for the home bias. Of course, home bias could arise if an investor anticipated holding lots of home assets: $q_h > q_f$. But then home bias comes not from processing costs, but from portfolio expectations. This is the mechanism explored in section 3.

The *shadow* price of foreign information required to explain home bias ($\psi = 7.3$) seems out of line with 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.[6] If producing home information required one analyst, and producing foreign information required one analyst and one translator, then the translator's salary would have to be 6.3 times the analyst's. Agency problems and legal/accounting differences could add costs, but the costs must be large. This explanation is also inconsistent with local bias and patterns of foreign investment (section 4.1, 4.2).

# 3 Main Results

The previous section showed that sustaining information asymmetry, without considering the interaction between learning and investment choices, is an uphill battle. This section analyzes a model where small asymmetries in investors' information not only persist, but are magnified. The only

[5]The investor's optimal portfolio is: $q^\star = 1/\rho\widehat{\Sigma}^{-1}(\hat{\mu} - pr)$. If home and foreign expected returns $(\hat{\mu} - pr)$ are equal and covariance is zero, then $q_h^\star/q_f^\star = \hat{\sigma}_f^2/\hat{\sigma}_h^2$.

[6]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.

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 backward induction, starting with the optimal portfolio decision, taking information choices as given. Given posterior mean $\hat{\mu}^j$ and variance $\widehat{\Sigma}^j$ of asset payoffs, the portfolio for investor $j$, from either country, is

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

Aggregating these asset demand across investors and imposing the market clearing condition delivers a solution for the equilibrium asset price level. Proof is in appendix B.3, along with the formulas for $A$ and $C$.

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

### 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 (8), into expected utility (7). Combining terms yields

$$U = E\left[\frac{1}{2}(\hat{\mu}^j - pr)'(\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr)|\mu, \Sigma\right]. \tag{9}$$

At time 1, $(\hat{\mu}^j - pr)$ is a normal variable, with mean $(-A)$ and variance $\Sigma_p - \widehat{\Sigma}^j$.[7] Thus, expected utility is the mean of a chi-square. Using the fact that the choice variable $\hat{\Lambda}$ is a diagonal matrix, that $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$, the formula for $A$ (equation 21), and the formula for the mean of a chi-square, we

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

can rewrite the period-1 objective as:

$$\max_{\hat{\Lambda}^j} \sum_i \left(\Lambda_{pi} + (\rho\Gamma_i'\bar{x}\hat{\Lambda}_i^a)^2\right)(\hat{\Lambda}_i^j)^{-1} \quad s.t.\ (3) \text{ and } (4) \tag{10}$$

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

The key feature of the learning problem (10) is its convexity in the posterior variance $(\hat{\Lambda}^j)$. In a 2-risk factor setting, the objective is $U = L_1/\hat{\Lambda}_1 + L_2/\hat{\Lambda}_2$, for positive scalars $L_1, L_2$. Thus, an indifference curve is $\hat{\Lambda}_2 = L_2\hat{\Lambda}_1/(U\hat{\Lambda}_1 - L_1)$, which asymptotes to $\infty$ at $\hat{\Lambda}_1 = L_1/U > 0$. The capacity constraint is $\hat{\Lambda}_2 = e^{2K}/\hat{\Lambda}_1$, which asymptotes to $\infty$ at $\hat{\Lambda}_1 = 0$. Because the indifference curve is always crossing the capacity constraint from below, the solution is always a corner solution. Figure 2 plots the indifference curve for a case where $L_1 = L_2$, the capacity constraint, and the no-negative learning bounds for our model (left panel) and the exogenous-portfolio model in section 2 (right panel).

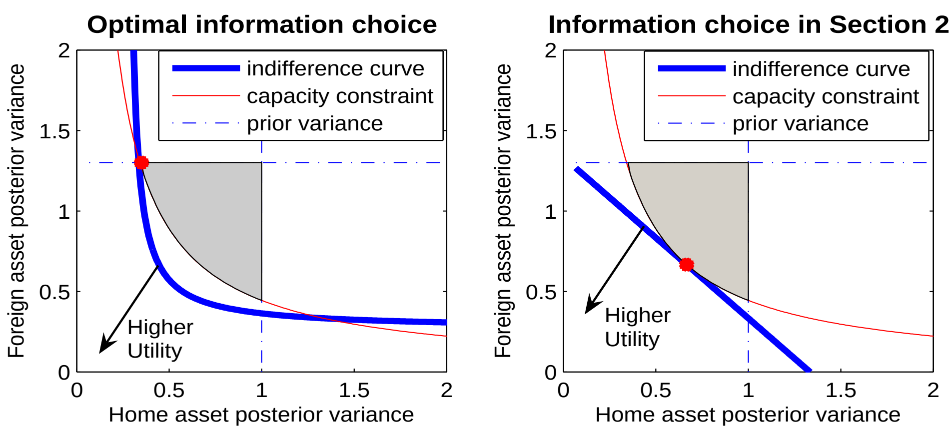


Figure 2: Objective and constraints in the optimal learning problem with 2 risk factors.

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

**Proposition 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 in appendix B.4. 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 $\Lambda_{pi}$), and higher returns: $\Gamma_i' E[f - pr] = \rho\hat{\Lambda}_i^a\Gamma_i'\bar{x}$. (See appendix B.3.) 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 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 $\Lambda$, 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 $\Lambda_{pi}$ 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. 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, $\Lambda_f$ represents home investors' prior uncertainty about foreign risk factors and $\hat{\Lambda}^\star_h$ represents foreigners' posterior uncertainty about home risks. Appendix B.5 proves the following two corollaries.

**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.*

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 $\Lambda^j_i$ is low. A low $\Lambda^j_i$ 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 $\frac{|\hat{\Lambda}_f|}{|\hat{\Lambda}^\star_f|} \geq \frac{|\Lambda_f|}{|\Lambda^\star_f|}$.*

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.

### 3.3 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 $\bar{x}$.

$$E[q^{no\ adv}] = \bar{x} \tag{11}$$

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 et al. (2004) capture when they estimate the home bias that uncertainty about foreign accounting standards could generate.

$$E[q^{no\ learn}] = \Gamma \Lambda^{-1} \Lambda^{a} \Gamma' \bar{x}, \tag{12}$$

where $\Lambda^a$ 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 (13)) 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:

$$E[q] = \Gamma\hat{\Lambda}^{-1}\hat{\Lambda}^a\Gamma'\bar{x} \tag{13}$$

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 first effect can be seen from equation (8): The entries of $\widehat{\Sigma}^{-1}$ corresponding to home risks are higher. 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 second effect, 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, raises $\hat{\Lambda}^a$ and thus the asset's return. Home investors are being compensated for more risk than they bear ($\hat{\Lambda}_i^a > \hat{\Lambda}_i^j$ in equation 13). Based on their information, this asset delivers high risk-adjusted returns. High returns make a long position optimal, on average. Both the magnitude and the general equilibrium effect increase home bias.[8]

The next two propositions (proven in appendix B.6) formalize the difference between the optimal portfolio (13), and the benchmark portfolios (12) and (11). Let $\Gamma_h$ be a sum of the eigenvectors in $\Gamma$ 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 $\Lambda_{hi} = \Lambda^\star_{fi}$ 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 ($K = 0$): $\Gamma_h' E[q] > \Gamma_h' E[q^{no\ learn}] > \Gamma_h' E[q^{no\ adv}]$.*

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

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

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 $\Lambda$, 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 toward 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 ($K = 0$): $\Gamma_h' E[q] \geq \Gamma_h' E[q^{no\ learn}] > \Gamma_h' E[q^{no\ adv}]$.*

When market size is different across countries, no home investor will learn more about foreign risks than any foreign investor will, and vice versa (corollary 3). Therefore, the average home investor still knows more about home risks, and tilts his portfolio to create home bias. When home risk factors are small, home investors are more likely to learn about larger foreign risks, and reduce their home bias. This prediction fits with cross-country data. 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). In the most extreme case, all investors learn about one risk. The ratio of home investors' and the average investor's posterior variance is the same whether investors learn or not: $E[q] = E[q^{no\ learn}]$. In this extreme case, learning does not amplify the home bias, but it doesn't undo it as it did in section 2.

The next proposition (proven in appendix B.7) shows that home investors earn higher returns on home assets. This is consistent with evidence found by Hau (2001). Following Admati (1985), we define the 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: $\partial E[(\Gamma_i' q)(\Gamma_i'(f - pr))]/\partial K > 0$.*

Most foreigners don't learn about home assets, but hold them as part of their diversified portfolio. Home investors use superior information on home assets to take advantage of foreigners. The more learning capacity the home investor has, the stronger the more profitable this strategy is.

It is not the case that an investor raises his return $(f_i - p_i r)$ 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 $(f - pr)$. He holds a long position in the asset when high returns are likely, and shorts the asset when very low returns are likely.

**The declining home bias** At first glance, these results seem to suggest that home bias should increase over time, as information becomes easier to acquire (capacity increases). This would be inconsistent with the modest decline in the U.S. home bias. However, that conclusion relies on home and foreign investors having the same capacity. If home investors' capacity increases more, returns on home assets decline. Relatively higher foreign returns may induce some home investors to specialize in learning about and holding foreign equity. Thus, asymmetric increases in capacity could reduce the average investor's home bias. Section 4.2 elaborates on this effect.

Furthermore, capital flow liberalization and increases in equity listings in the last 30 years has created more investible foreign risk factors. Model investors would add these assets to the 'no advantage' part of their portfolio ($\bar{x}$). This effect would contribute to the increase in foreign equity investment, and also reduce home bias.

### 3.4 A Numerical Example

A numerical example quantifies the potential home bias in the model. We start from the symmetric setup of proposition 3. There are two countries, home and foreign, with 1000 investors in each. Risk aversion is $\rho = 2$. There are 5 home and 5 foreign assets. Information advantages are symmetric: Foreigners start out $\alpha$ times more uncertain about home risks $(1+\alpha)\Lambda_h = \Lambda_h^\star$, and home investors are $\alpha$ times more uncertain about foreign risks $\Lambda_f = (1+\alpha)\Lambda_f^\star$. We consider a 10% information advantage ($\alpha = 0.1$). The supply of each asset has mean ($\bar{x}$ =100) and standard deviation 10. Expected payoffs for home and foreign assets are equal. They are equally spaced between 1 and 2. The mean of the average investor's prior belief is equal to true asset payoffs. The standard deviation of prior beliefs is between 15-30%, such that all assets have the same expected payoff to standard deviation ratio. We vary learning capacity $K$ to explore its effect.

Following convention, we define the home bias as

$$\text{home bias} = 1 - \frac{1 - share\ of\ home\ asset\ in\ home\ portfolio}{share\ of\ foreign\ assets\ in\ world\ portfolio} \tag{14}$$

In this example, the share of foreign assets in the world portfolio is 0.5. We compare the home bias our model generates to two benchmarks. 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). The ten percent initial information advantage by itself generates a 5.3 percent home bias in asset holdings.

**Uncorrelated assets** In our model, learning capacity magnifies home bias because home investors specialize in learning about one home asset. When there is only enough capacity to view a signal that eliminates 22 percent of the risk in one asset (K=.24, $1 - e^{-K} = .22$), the home bias almost doubles from 5.3 to 10 percent. When there is enough capacity to eliminate 70 percent of the risk in one asset (K=1.2, $1 - e^{-K} = .70$), the home bias is 45%, more than eight times larger than the home bias without learning.

**Correlated Home Assets** When home assets are positively correlated with each other, and foreign assets are positively correlated with each other (correlations of 10-30%), but the two sets of assets are mutually uncorrelated, each investor learns about one risk factor, that all his domestic assets load on. Introducing moderate correlation strongly increases the home bias. 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. Home bias doubles to 19.4% when $1 - e^{-K} = .22$; it increases to 59.5% for $1 - e^{-K} = .70$. (See line with circles in figure 3.) In contrast, the no learning benchmark is unaffected (5.3%, line with diamonds). With $1 - e^{-K} = .82$, home bias is 72%, just shy of the 76% observed in the data. In the next section, we endow investors with advantages in local risks, rather than country risks. This change reduces the capacity required to match the home bias.

# 4 Extensions

## 4.1 Local Information: An Amplification Mechanism

A unified explanation for home and local bias is something that many home bias theories cannot provide (Coval and Moskowitz 2001). Their coexistence makes an information-based explanation more appealing than explanations rooted in exchange rate risk, institutional difference, or language barriers. By giving investors slightly more precise signals over local assets, this model can explain the local investment bias, and the accompanying local excess returns. We could also interpret this

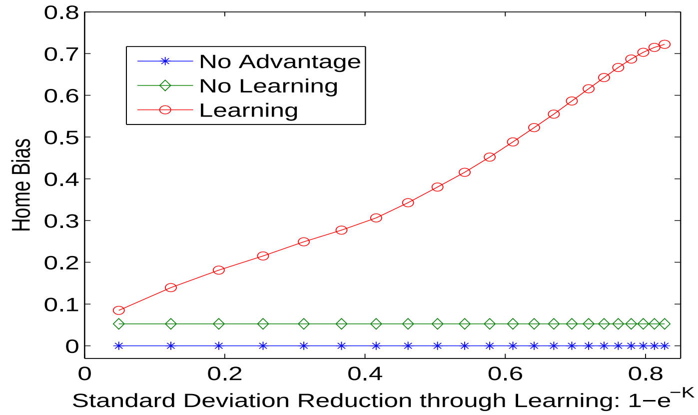


Figure 3: Home Bias Increases With Capacity. Numerical example with two countries. Assets within a country have correlated payoffs (cov= $.09^2$). Home bias is defined in (14). The 'no advantage' line (stars) gives the home bias in an economy with no initial informational advantage and no capacity to 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. Learning capacity K varies between 0.05 and 1.5. The horizontal axis plots the potential percentage reduction in the standard deviation of one asset.

as a model of bias toward industries that investors have prior knowledge of, perhaps because of their job.

Suppose that home investors each had an advantage in only one home risk factor, the one most concentrated in their region's asset. An investor $j$ from region $m$ draws an independent prior belief from the distribution $\mu^j \sim N(f, \Sigma_m)$, where $\Sigma_m = \Gamma \Lambda_m \Gamma$, and $\Lambda_m$ has a $m$th diagonal entry that was lower than the $m$th diagonal in the beliefs of any other region. In this model, local investors have an incentive to learn more about their local assets, because of their initial information advantage (proposition 3). As a result, portfolios weight local assets more heavily, and earn higher profits from on local assets, on average (proposition 6). Local advantages also *amplify the effects* of home advantages: When fewer investors share an advantage in the same local risk, locals have a larger advantage relative to the average investor (higher $\hat{\Lambda}^a_m / \Lambda^j_m$). A more specialized advantage magnifies the optimal portfolio bias ($E[\Gamma'_m q = \hat{\Lambda}^a_m / \Lambda^j_m (\Gamma'_m \bar{x}))$. Because returns to specialization increase when information advantages are more concentrated, investors diversify less.

**Numerical Example** To quantify local bias, we use the same numerical example as in section 3.4, with correlated assets. We measure local bias as in (14), treating localities like countries. The only difference is that instead of giving 1000 home (foreign) investors a 10% initial information advantage in all 5 home (foreign) assets, we give 200 investors each a 50% information advantage in one local asset; the aggregate information advantages at home and abroad are unchanged. Without learning, the average local bias is only 5.0%; with capacity, $1 - e^{-K} = 0.70$, it is 30%. The average local investor holds 3.6 times what a diversified investor would hold, of his local asset.

Concentrating information advantages in local assets *increases home bias.* Without learning, the home bias is 8%; with low capacity ($1-e^{-K} = 0.22$), it is 23%. With more capacity ($1-e^{-K} = 0.70$), home bias is 76%, which is 29% more than the home advantage learning model. This matches the 76% home bias in the data.

**How large is capacity?** One way of inferring capacity is from excess profits that locally-concentrated investors earn. Using brokerage account data, Ivkovic et al. (2005) show that individuals investors with concentrated portfolios earn 10% higher risk-adjusted annual returns on local non-S&P500 stocks than investors with diversified portfolios. In our simulated model, we compute the analogous risk-adjusted return statistics, to determine if our capacity level is realistic.

To link the model to data, we equate the largest risk-factor in each country (80% of market capitalization) with S&P500 stocks (73% of US market capitalization). We compare expected returns of investors who learn about their local asset and non-local investors on the non-S&P500 risk factors. Because informed local investors hold long positions only when assets are likely to have high payoffs, they earn higher average returns. For the level of capacity that matches the empirical home bias ($1 - e^{-K} = .70$), local investors' return on the smaller risk factors is 5% higher than what non-locals earn. The model can match Ivkovic et al. (2005)'s 10% result for ($1 - e^{-K} = .75$). These results suggest that the level of capacity required to match the home bias is not outlandish; it may even underestimate investors' true capacity.

Ivkovic et al. (2005) focus on non-S&P500 stocks because their informational asymmetries are potentially the largest. They report insignificant outperformance on the S&P assets. While our model cannot speak to the statistical significance of their results, it does qualitatively match the pattern of lower outperformance on larger assets. For the calibration that matches the home bias, local investors' return on the S&P risk factors is only 2% higher than what non-locals earn. Returns fall on the large risks because their size makes them valuable to learn about. Low average uncertainty about the risks makes equilibrium returns and outperformance fall.

## 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 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 some 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 and hold Taiwanese assets, 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 investment risk because of their initial information disadvantage. 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. If Americans have capacity that exceeds Taiwanese capacity, by more than their initial advantage in some risk factor, then Americans can become better informed about that risk. By proposition 6, being more informed than the average Taiwanese investor implies than the American investor will out-perform the average investor in assets that load on his researched risk factor ($fi$).[9]

The risk factors that Americans should learn about, hold more of, and profit from are ones that the largest assets weight most heavily on. Assuming that the average uncertainty ($\hat{\Lambda}^a$), noise in prices ($\Lambda_p$) and American uncertainty ($\Lambda$) 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). 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.

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

## 5 Conclusions

This paper questions the common assumption that residents have more information about their region's assets than do non-residents. The question is: If investors are restricted in the amount of information they can learn about risky asset payoffs, which assets would they choose to learn about? 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 informational asymmetries. An initial information advantage in a risk is like a comparative advantage in learning about that risk. Investors believe that they will hold slightly more of the asset they have an advantage in than the average investor will. Knowing that the asset will be a larger part of their portfolio causes investors to value learning about it more. By each specializing 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 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. 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 Stulz (1997).

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 is an example of a local information advantage. Such an advantage can lead an investor to rationally learn more about his own country, own industry, and own company assets, and to weight them heavily in his portfolio.[10] Massa and Simonov (2005) document that Swedish investors tilt their portfolio towards the industry they work in, and conclude that this is a rational response to

[10]See Van Nieuwerburgh and Veldkamp (2006) for an analysis of the role of information choice in the own-company stock puzzle.

being better informed.

Future work should focus on building a dynamic model of learning and investing to explain the patterns of equity flows. The large cross-border equity flows observed have proven difficult to reconcile with many theories of home bias (Tesar and Werner 1995). Gravity model estimates (Portes, Rey and Oh 2001) add to the puzzle by showing that investors trade most with countries whose assets offer little diversification benefit. Perhaps it is the fact that these assets are poor diversification devices, that makes it efficient to learn about them, to hold them, and to trade them prolifically. Direct evidence on information flows support this explanation. Portes and Rey (2003) find that nearby markets exchange abundant information: they exhibit high telephone traffic and strong evidence of insider trading.

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

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

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

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

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

**Baxter, Marianne and Urban Jermann**, "The International Diversification Puzzle Is Worse Than You Think," *The American Economic Review*, 1997, *87(1)*, 179–180.

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

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

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

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

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

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

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

**Dvorak, Tomas**, "Do Domestic Investors Have an Information Advantage? Evidence from Indonesia," *The Journal of Finance*, 2005, *60*, 817–839.

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

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

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

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

**Hatchondo, Juan Carlos**, "Asymmetric Information and the Lack of International Portfolio Diversification," 2004. Working Paper.

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

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

**Ivkovic, Zoran, Clemens Sialm, and Scott Weisbenner**, "Portfolio Concentration and the Performance of Individual Investors," 2005. Working Paper.

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

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

**Massa, Massimo and Andrei Simonov**, "Hedging, Familiarity and Portfolio Choice," *The Review of Financial Studies*, 2005, *Forthcoming.*

**Mondria, Jordi**, "Financial Contagion and Attention Allocation," 2005. Working Paper.

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

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

**Portes, Richard and Helene Rey**, "The Determinants of Cross-Border Equity Flows," 2003. Working Paper.

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

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

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

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

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

**Tesar, Linda and Ingrid Werner**, "Home Bias and High Turnover," *Journal of International Money and Finance*, 1995, *14(4)*, 467–92.

**____ and ____** , "The Internationalization of Securities Markets Since the 1987 Crash," 1998. Brookings-Wharton Papers on Financial Services, The Brookings Institution.

**Van Nieuwerburgh, Stijn and Laura Veldkamp**, "Information Acquisition and Portfolio Under-Diversification," 2005. Working Paper.

—— **and** ——, "Inside Information and the Own Company Stock Puzzle," *Journal of the European Economic Association, Papers and Proceedings*, 2006, *May*, forthcoming.

# A The Model: Technical Details

## A.1 The Role of Preferences

Acquiring information does not change asset fundamentals. In particular, it cannot change the total amount of uncertainty about the assets' payoffs. All it does is cause that uncertainty to be resolved sooner. It reduces the risk that an investor faces at the interim stage, after learning but before investing.

Because standard expected utility delivers no preference for early resolution of uncertainty, we use a utility function, similar to Epstein-Zin preferences, that makes early resolution desirable. This formulation keeps the exponential structure of preferences that makes the problem tractable. It also treats learned information and prior information as equivalent. We are asking what variance of beliefs our investors would most like to have, in period 2, when they decide how to invest.

If we instead use standard expected utility, the learning choice is indeterminate. Capacity increases expected utility linearly, but how it is allocated makes no difference. In a sense, this is already a success because the strict preference for diversification dictated by standard theory has been undone. But because there would be no testable implications, this does not make for a satisfying theory. The indeterminacy can be resolved by introducing a small preference for early resolution of uncertainty.

Using these preferences is not in any way assuming the result. The fixed-portfolio model in section 2 generates the opposite conclusion, despite having the same preferences. That model also illustrates why using preferences for early resolution of uncertainty is the most sensible way to analyze learning choices. Without a preference for early resolution of uncertainty, the investor who takes his portfolio as fixed when he chooses what to learn wouldn't care what he learned about, or if he learned at all. With a preference for late resolution of uncertainty, the fixed-portfolio investor would be averse to observing new information (ignorance is bliss). In order to meaningfully understand the information choices investors make, we start not with investors who prefer ignorance, but with investors who prefer to learn.

Of course, one could break indifference by changing the learning technology, rather than changing the preferences. But then agents would specialize or diversify simply because the learning technology dictated that such a strategy yielded a higher return. With the technology in (3), investors get the same expected return, no matter how capacity is allocated. In other words, capacity $K$ is a sufficient statistic for expected returns. That is the sense in which *our is the unique, expected-returns-neutral learning technology.* Keeping this technology and assuming aversion to risk at the time of investment, assures us that the results come only from investors' desire to efficiently reduce risk. They are not driven by the technology itself.

## A.2 Implications of the Learning Technology

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. In this appendix, we show that neither force depends on the assumption that investors learn about risk factors that are principal components.

In contrast, Mondria (2005) allows agents to learn about arbitrary risk factors, but restricts his analysis to symmetric equilibria, where all investors start with the same information and learn the same information. His results do not overturn ours. In fact, they demonstrate the robustness of our main result to changes in the learning technology. When an investor in Mondria's model has an initial advantage in an asset, that investor will also learn more about that asset. Thus specialization arises, even with risk factor choice. In both models, equilibria where all investors learn the same information do exist, but only when investors start with identical prior beliefs (no initial advantages) and have low capacity.

**Specialization persists with arbitrary risk factors** Let the posterior covariance matrix be $\widehat{\Sigma} = \widehat{\Gamma}'\hat{\Lambda}\widehat{\Gamma}$, where $\widehat{\Gamma}$ is an *arbitrary* posterior eigenvector matrix of $\widehat{\Sigma}$. This relaxes the constraint that signals must be about principal components of payoffs. Imposing the existence of some eigenvectors for $\widehat{\Sigma}$ is not restrictive because every variance-covariance matrix is positive-definite and has an eigen-representation of this form. The objective function (10) then becomes:

$$\max_{\hat{\Lambda}} \frac{1}{2}Tr\left(\widehat{\Gamma}\hat{\Lambda}^{-1}\widehat{\Gamma}'\Sigma_p - I\right) + \frac{1}{2}E[f-pr]'\widehat{\Gamma}\hat{\Lambda}^{-1}\widehat{\Gamma}'E[f-pr]. \tag{15}$$

Note that this objective is still a linear function of the form $U = \alpha + \sum_i \hat{L}_i\hat{\Lambda}_i^{-1}$, for positive constants $\alpha$ and $\{\hat{L}_i\}_i$. The learning constraints (3) and (4) are still inequality and product constraints on the eigenvalues of posterior beliefs $\hat{\Lambda}^{-1}$. Since the problem takes the same form of maximizing a sum, subject to a product constraint, specialization will still arise. (See appendix B.4.) Furthermore, the new learning index is $\hat{L}_i\Lambda_i^{-1}$, which is increasing in prior precision $\Lambda_i^{-1}$. Thus, investors prefer to learn about risks that they already know well.

In fact, the specialization result also arises in Mondria (proposition 3). His agents devote all their learning capacity to one linear combination of asset payoffs, as do ours. The differences in our results are purely about whether investors want to specialize in the same risks, or different ones. Next, we show that investors still prefer to learn about different risks.

**Strategic substitutability persists with arbitrary risk factors** Choosing risk factors does not eliminate the strategic substitutability. Risks that other investors learn about will still have lower returns. That is what makes one investor want to make her information set as different as possible from the other investors'.

First, we introduce new 'aggregate risk factors'. Note that Admati's (1985) proof of the equilibrium price does not depend on any particular risk structure. The key information variables in the price formulae, (18), (19) and (20), will have the same forms as in appendix B.3, but with the eigenvectors of the average investor's posterior beliefs $\tilde{\Gamma}$, s.t. $\Sigma_\eta^a = \tilde{\Gamma}\Lambda_\eta^a\tilde{\Gamma}'$. The columns $\tilde{\Gamma}_i$ are 'aggregate risks' and no longer have to be the same as the eigenvectors of prior beliefs ($\tilde{\Gamma} \neq \Gamma$) or the posterior risk factors of any given investor ($\tilde{\Gamma} \neq \widehat{\Gamma}$).

An individual's value of learning about a risk $\widehat{\Gamma}_i$ correlated with $\tilde{\Gamma}_j$ is decreasing in the precision of the average investor's beliefs about the payoffs of aggregate risk $j$. The average investor's information enters an individual's utility function (15) in two places: through $\Sigma_p$ in the first term and through $E[f-pr]$ in the second term.

Define a weighting matrix $M$ such that $\widehat{\Gamma} = \tilde{\Gamma}M$. Then $M_{ij}$ describes how much the individual's risk factor $j$ weights on the aggregate risk factor $i$, and thus how much their payoffs covary. Let $M_i$ be the $i$th row of $M$.

The first term in utility is $\frac{1}{2}Tr\left(\widehat{\Gamma}'\hat{\Lambda}^{-1}\widehat{\Gamma}\Sigma_p - I\right)$. Then, the marginal utility of increasing the precision of beliefs about private risk $i$, $\hat{\Lambda}_{ii}^{-1}$ is $\frac{1}{2}Tr(M_i'M_i\Lambda_p)$. The partial derivative of this marginal utility with respect to $\Lambda_{pjj}$ is the $(j,j)$th entry of $(1/2M_i'M_i)$, which is $1/2(M_{ij})^2 = 1/2(\tilde{\Gamma}_j'\widehat{\Gamma}_i)^2$, the squared covariance of private risk $i$ with aggregate risk $j$. Since the squared covariance is always positive, the value of learning is increasing the in the exploitable pricing error. The last step links this to the average investor's uncertainty: $\partial\Lambda_{pjj}/\partial\Lambda_{\eta j}^a > 0$ because equation (19) tells us that $\Lambda_{pjj} = (\Lambda_\eta^a)^2$, times a positive constant, and $\Lambda_\eta^a$ is always positive. Thus, the first component of the value of learning about $\widehat{\Gamma}_i$ is always weakly increasing in aggregate uncertainty $\Lambda_{\eta j}^a$. This substitutability is stronger, the more $\widehat{\Gamma}_i$ and $\tilde{\Gamma}_j$ covary.

The second term can be rewritten as $\sum_i(\widehat{\Gamma}_i'E[f-pr])^2\hat{\Lambda}_i^{-1}$. The marginal utility of increasing the precision of beliefs about private risk $i$ is $(\widehat{\Gamma}_i'E[f-pr])^2$. The expected return on risk factor $\widehat{\Gamma}_i$ is a weighted sum of the expected returns on each risk $\tilde{\Gamma}_j$: $\widehat{\Gamma}_i'E[f-pr] = \sum_j M_{ij}\tilde{\Gamma}_j'E[f-pr]$. Manipulating equation (16) reveals the expected return on aggregate risk $j$: $\tilde{\Gamma}_j'E[f-pr] = \rho[\frac{1}{\rho^2\sigma_x^2}(\Lambda_{\eta j}^a)^{-2} + (\Lambda_{\eta j}^a)^{-1}]^{-1}\tilde{\Gamma}_j'\bar{x}$. This converges to zero as the average investor's precision of beliefs $(\Lambda_{\eta j}^a)^{-1}$ rises. Thus the utility gain $(\widehat{\Gamma}_i'E[f-pr])^2$ from learning about the payoff of $\widehat{\Gamma}_i$ decreases in the aggregate amount learned about each of its aggregate risk components because the expected return on each $\tilde{\Gamma}_j$ converges to zero as $(\Lambda_{\eta j}^a)^{-1}$ increases.

In other words, no matter what risk factors we learn about, when others learn about correlated risks, it reduces the expected returns and exploitable pricing errors on our risk factors and makes them less desirable for to learn about. This is strategic substitutability.

# B Proofs

## B.1 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 $\upsilon$ is the Lagrange multiplier on the capacity constraint and $\phi_i$ 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 $\hat{\Lambda}_i > \Lambda_i$. □

## B.2 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 $\epsilon$ 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

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

□

## B.3 Proof of Proposition 2

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

$$A = -\rho \left(\frac{1}{\rho^2 \sigma_x^2} (\Sigma_\eta^a \Sigma_\eta^{a'})^{-1} + (\Sigma_\eta^a)^{-1}\right)^{-1} \bar{x}, \quad (16)$$

$$C = -\left(\frac{1}{\rho^2 \sigma_x^2} (\Sigma_\eta^a \Sigma_\eta^{a'})^{-1} + (\Sigma_\eta^a)^{-1}\right)^{-1} \left(\rho I + \frac{1}{\rho \sigma_x^2} (\Sigma_\eta^a)^{-1'}\right). \quad (17)$$

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.

$$
\begin{aligned}
(\Sigma^a_\eta)^{-1} &= \Gamma(\Lambda^a_\eta)^{-1}\Gamma' = \frac{1}{2}\Sigma^{-1} + \frac{1}{2}(\Sigma^\star)^{-1} + \int_j (\Sigma^j_\eta)^{-1} dj, && (18)\\
(\Sigma_p)^{-1} &= \Gamma\Lambda_p^{-1}\Gamma' = \frac{1}{\rho^2\sigma_x^2}(\Sigma^a_\eta\Sigma^{a'}_\eta)^{-1}, && (19)\\
(\widehat{\Sigma}^a)^{-1} &= \Gamma\hat{\Lambda}_a^{-1}\Gamma' = \frac{1}{\rho^2\sigma_x^2}(\Sigma^a_\eta\Sigma^{a'}_\eta)^{-1} + (\Sigma^a_\eta)^{-1} && (20)
\end{aligned}
$$

We have assumed that investors choose to obtain signals about the eigenvectors $\Gamma$ of the prior covariance matrix $\Sigma$. It is easy to show that when $\Sigma_\eta$ has eigenvectors $\Gamma$, the three precision matrices above also have the same eigenvectors.

We note and later use that $CC'\sigma_x^2 = \rho^2\sigma_x^2\Sigma^a_\eta\Sigma^{a\prime}_\eta = \Sigma_p$, because $C = -\rho\Sigma^a_\eta$. We also use that

$$
-\Gamma'_i A = \rho\Gamma'_i\hat{\Sigma}^a\bar{x} = \rho\Gamma'_i\Gamma\hat{\Lambda}^a\Gamma'\bar{x} = \rho(\Gamma'_i\bar{x})\hat{\Lambda}^a_i, \tag{21}
$$

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 $\Gamma'\Gamma = I$. □

## B.4 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 $y_iX_i$, and $Tr\left(\widehat{\Sigma}^{-1}\Sigma\Sigma^{-1}(V_{ER} + \widehat{\Sigma})\right) = \sum_{i=1}^N X_iy_i$. This is because $\Sigma$, $\widehat{\Sigma}$, and $C$ all share the same eigenvectors $\Gamma$. 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 $\tau = e^{2K}\left(\prod_{j=1}^N\left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right)\right)^{-1}$. We endow the investor with enough capacity such that he has spare capacity to acquire private signals after devoting capacity to learning from prices: $\prod_{j=1}^N\left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right) < e^{2K}$ and therefore $\tau > 1$. For future reference define the 'spare capacity' of an investor who learns about risk factor $i$ as

$$
\tilde{K}_i = K - \frac{1}{2}\sum_{j\neq i}\log\left(1 + \frac{\Lambda_{pj}^{-1}}{\Lambda_j^{-1}}\right). \tag{22}
$$

□

## B.5 Proof of Corollaries 2 and 3

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

$$
\frac{\Lambda_{pi}}{\Lambda_i} + \frac{(\hat{\Lambda}^a_i)^2}{\Lambda_i}(\Gamma'_i\bar{x})^2 > \frac{\Lambda_{pi}}{\Lambda^\star_i} + \frac{(\hat{\Lambda}^a_i)^2}{\Lambda^\star_i}(\Gamma'_i\bar{x})^2. \tag{23}
$$

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

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

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 (22). 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 ($\Lambda_p^{-1}$) 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 $j$:

$$\frac{\Lambda_{pi}}{\Lambda_i} + \frac{(\hat{\Lambda}_i^a)^2}{\Lambda_i}(\Gamma_i'\bar{x})^2 = \frac{\Lambda_{pj}}{\Lambda_j^\star} + \frac{(\hat{\Lambda}_j^a)^2}{\Lambda_j^\star}(\Gamma_j'\bar{x})^2. \tag{25}$$

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 $(\hat{\Lambda}_a)^2/\Lambda$ and $\Lambda_p$ 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. $\Box$

## B.6 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 $\xi_i$ 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 (13), describing the portfolio with $K > 0$ and equation (12), 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 $\xi_i$ ($\xi_i^\star$) 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 (13) by $\Gamma_i'$. 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: $E[q_i] > E[q_i^{no\ learn}]$.$\square$

## B.7 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:

$$\text{Expected return } = (\Gamma_i' E[q])(\Gamma_i' E[(f - pr)]) + \Gamma_i' cov(q, f - pr)\Gamma_i. \tag{26}$$

Using equation (13) and the price formula in appendix B.3,it can be shown that $cov(q, f - pr) = \frac{1}{\rho}(\widehat{\Sigma}^{-1}\Sigma - I)$. Canceling out orthogonal eigenvectors, we can rewrite:

$$\text{Expected return } = \frac{1}{\rho}\left(\hat{\Lambda}_i^{-1}(\Gamma_i' A)^2 + \hat{\Lambda}_i^{-1}\Lambda_i - 1\right). \tag{27}$$

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 $\Gamma_i$ or $\Lambda_i$. Therefore, capacity increases the expected return. $\square$