Internet Appendix for "Information Immobility and the Home Bias Puzzle" *

This technical appendix comprises three sections. The first provides step-by-step mathematical guidance on how to solve the model. It derives expected utility, expected portfolio holdings, the strategic substitutability and equilibrium uniqueness results, and the cost of foreign information in the model without increasing returns. The second section explores the model's assumptions. In particular, it looks at the foundations for expected mean-variance preferences, it allows investors to choose how much to learn, and it considers investors who can learn about risk factors other than principal components. The third section derives the formal links between the theory's predictions and the empirical evidence we offer in its support. Exploring these links also reveals why some findings that purportedly refute the information-based explanation for underdiversification are not, in fact, inconsistent with the predictions of a rational information-based theory.

A. Deriving the Model's Main Results

A.1. Taking the Expectation of the Utility Function

Recall that expected utility (equation (7) in the main text) is given by

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

At time 1, $(\hat{\mu}^j - pr)$ is a normal variable, with mean $(-A) = \rho\widehat{\Sigma}^a\bar{x}$ (equation (13) in the main text) and variance $\Sigma_p - \widehat{\Sigma}^j$. To derive this variance, note that $\text{var}(\hat{\mu}|\mu) = \Sigma - \widehat{\Sigma}$, $\text{var}(pr|\mu) = \Sigma + \Sigma_p$, and $\text{cov}(\hat{\mu}, pr) = \Sigma$.

Thus, expected utility is the mean of a noncentral χ^2. If a generic random variable z is normally distributed $z \sim N(\mu, \Sigma)$, then $E[z'z] = trace(\Sigma) + \mu'\mu$. Applying this formula yields

$$U = \left\{\frac{1}{2}Tr\left(\widehat{\Sigma}^{-1}(\Sigma_p - \widehat{\Sigma})\right) + \frac{1}{2}A'\widehat{\Sigma}^{-1}A\right\}. \tag{IA.2}$$

Since $\widehat{\Sigma}$ and Σ_p have the same eigenvectors Γ, we can rewrite the trace term as $Tr\left(\widehat{\Sigma}^{-1}(\Sigma_p - \widehat{\Sigma})\right) = Tr\left(\Gamma\hat{\Lambda}^{-1}\Gamma'\Gamma(\Lambda_p - \hat{\Lambda})\Gamma'\right)$. Recall that the trace of an $N \times N$ matrix is the sum of its N eigenvalues. Therefore, the trace term is equal to $\sum_i^N \hat{\Lambda}_i^{-1}(\Lambda_{pi} - \hat{\Lambda}_i)$.

*Please cite this supplement document as: Van Nieuwerburgh, Stijn and Laura Veldkamp, 2008, Internet Appendix for "Information Immobility and the Home Bias Puzzle," *Journal of Finance* 64, 1187-1215, `http://www.afajof.org/IA/2009.asp`. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

The second term in equation (IA.2) above can be rewritten as well. Substituting in for A yields $\rho^2\bar{x}'\widehat{\Sigma}^a\widehat{\Sigma}^{-1}\widehat{\Sigma}^a\bar{x}$. Rewriting the variance matrices in terms of the eigen-decomposition results in $\rho^2(\Gamma'\bar{x})'(\hat{\Lambda}^a)^2\hat{\Lambda}^{-1}\Gamma'\bar{x}$. Note that $\Gamma'\bar{x}$ is a vector and $(\hat{\Lambda}^a)^2\hat{\Lambda}^{-1}$ is a diagonal matrix. Therefore, this matrix product can be written as a sum of the products, entry-by-entry. The resulting expression for expected utility is

$$U = \frac{1}{2}\sum_i(\hat{\Lambda}_i^{-1}\Lambda_{pi} - 1) + \frac{\rho^2}{2}\sum_i(\Gamma_i'\bar{x})^2(\hat{\Lambda}_i^a)^2\hat{\Lambda}_i^{-1}.$$

Rearranging this expression and dropping the constant delivers the utility function in equation (9) of the main text.

A.2. Strategic Substitutability

Strategic substitutability in learning means that the value of learning about a risk factor declines as others learn more about it. Since the value of learning about a risk factor i to an investor j is summarized by the learning index $\mathcal{L}_i^j$ and the amount others know about that risk factor i is summarized by the average precision matrix $(\hat{\Lambda}_i^a)^{-1}$, strategic substitutability means that

$$\frac{\partial\mathcal{L}_i^j}{\partial(\hat{\Lambda}_i^a)^{-1}} < 0.$$

Substituting the definitions of $\hat{\Lambda}^a$ and Λ_p from equations (A5) and (A6) of the main text into the definition of the learning index yields

$$\mathcal{L}_i^j = \frac{(\Lambda_i^j)^{-1} + \Lambda_{\eta i}^{-2}/(\rho^2\sigma_x^2)}{(\Lambda_{\eta i}^{-1} + \Lambda_{\eta i}^{-2}/(\rho^2\sigma_x^2))^2}\rho^2(\Gamma_i'\bar{x})^2 + \rho^2\sigma_x^2\frac{(\Lambda_{\eta i}^{-2})^{-1}}{\Lambda_i^j}.$$

Taking a partial derivative reveals that $\partial\mathcal{L}_i^j/\partial\Lambda_{\eta i}^{-1} < 0$.

Taking a partial derivative of equation (A6) in the text reveals that $\partial(\hat{\Lambda}_i^a)^{-1}/\partial\Lambda_{\eta i}^{-1} > 0$. Applying the chain rule tells us that $\partial\Lambda_{\eta i}^{-1}/\partial(\hat{\Lambda}_i^a)^{-1} > 0$ as well. Applying the chain rule once more tells us that $\partial\mathcal{L}_i^j/\partial(\hat{\Lambda}_i^a)^{-1} = \partial\mathcal{L}_i^j/\partial\Lambda_{\eta i}^{-1}\cdot\partial\Lambda_{\eta i}^{-1}/\partial(\hat{\Lambda}_i^a)^{-1} < 0$.

Aggregate allocation of capacity is unique: Suppose that a mass of investors deviates from the equilibrium allocation and learns more about some risk k. If k is not in the investors' indifference set, then it is not an optimal information choice. Learning about some other risk with a higher index delivers greater expected utility. If k is the only risk in the investors' indifference set, then learning about it is the unique equilibrium strategy; this is not a deviation. If k was one of multiple risks in the investors' indifference set, then the additional mass of investors learning about it would lower $\mathcal{L}_k$, below the level of the other risk factors in the indifference set. Thus, learning about k would no long be optimal.

A.3. Expected Portfolio Holdings

The first-order condition for investment choice reveals that the optimal portfolio, conditional on an investor j's posterior beliefs $\hat{\mu}$ and $\widehat{\Sigma}$, is

$$q^j = \frac{1}{\rho}(\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr). \tag{IA.3}$$

This equation is the chosen portfolio at time 2. We want to know the expected portfolio at time 1, based only on prior information. Since signals are chosen based on prior information, the investor knows what his posterior variance $\widehat{\Sigma}^j$ will be. The investor needs to take an expectation over $(\hat{\mu}^j - pr)$. The fact that Bayesian beliefs are unbiased implies that $E[\hat{\mu}^j - pr] = E[f - pr]$. From Appendix C in the paper, we know that prices take the form $pr = A + f + Cx$, where x is a mean-zero random asset supply. Therefore, $E[f - pr] = -A$, where $A = -\rho\widehat{\Sigma}^a\bar{x}$ (equation (A2) in the main text). Substituting this in to (IA.3) yields

$$E[q^j] = \frac{1}{\rho}(\widehat{\Sigma}^j)^{-1}\rho\widehat{\Sigma}^a\bar{x} = (\widehat{\Sigma}^j)^{-1}\widehat{\Sigma}^a\bar{x}.$$

Replacing $\widehat{\Sigma}^j$ and $\widehat{\Sigma}^a$ with their eigen-decompositions, and using the facts that $\Gamma^{-1} = \Gamma'$ and $\Gamma'\Gamma = I$, delivers

$$E[q^j] = \Gamma(\hat{\Lambda}^j)^{-1}\hat{\Lambda}^a\Gamma'\bar{x}, \tag{IA.4}$$

which is the portfolio expression $E[q^\star]$ that appears in equation (10) of the main text. When investors have no capacity to learn ($K = 1$), then $\hat{\Lambda}^j = \Lambda^j$ and we obtain the no-learning portfolio $E[q^{no\,learn}]$. When investors have no capacity to learn ($K = 1$) and there are no initial information differences, then $\hat{\Lambda}^j = \Lambda^j = \hat{\Lambda}^a = \Lambda^a$ and we obtain the world market portfolio $E[q^{div}]$.

A.4. The Required Price of Foreign Information without Increasing Returns

This section considers the model of Section II in the main text, where investors choose what to learn, taking their expected portfolio holdings as given. Specifically, each investor expects to hold an equal amount of home and foreign assets ($q_h = q_f$). We ask the question: How much higher must the processing cost for foreign information be in order to quantitatively explain the degree of home bias in the data?

Suppose there is one home and one foreign asset, with prior variances σ_h^2 and σ_f^2, posterior variances $\hat{\sigma}_h^2$ and $\hat{\sigma}_f^2$, identical expected returns $(\hat{\mu}_h - p_h r)$ and $(\hat{\mu}_f - p_f r)$, and zero covariance. Replace the original capacity constraint (3) with a capacity constraint that requires ψ times more capacity to process foreign than home information: $(\sigma_h/\hat{\sigma}_h) \cdot (\sigma_f/\hat{\sigma}_f)^\psi \leq K$.

The optimal learning choice is described by the first-order conditions for $\hat{\sigma}_h^2$ and $\hat{\sigma}_f^2$. Combining the two first-order conditions tells us that, 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 optimal learning choice implies the following optimal portfolio holdings: $q_h^\star = \rho^{-1}\sigma_h^{-2}(\hat{\mu}_h - p_h r)$ and $q_f^\star = \rho^{-1}\sigma_f^{-2}(\hat{\mu}_f - p_f r)$. An investor in this world who chooses a portfolio for which $q_h^\star/q_f^\star = 7.3$, as the average U.S. investor in the data does, has beliefs that satisfy $q_h^\star/q_f^\star = \sigma_h^{-2}/\sigma_f^{-2} = 7.3$. From the first-order conditions for the optimal learning choice $\hat{\sigma}_f^2/\hat{\sigma}_h^2 = \psi q_h^2/q_f^2$ and the fact that $q_h = q_f$, we see that such an investor must have $\psi = 7.3$. This implies that the observed home bias can only be explained by the model of Section II when processing foreign information takes 7.3 times more capacity than processing the same amount of home information.

The zero covariance assumption in the above example biases ψ downward by overestimating home bias in two ways. First, it makes gains to diversification large. Second, if home signals are informative about correlated foreign assets, home bias would fall. For both reasons, the required cost differential would have to be even higher in the realistic case with positive covariance. 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 the model of Section III.

B. The Model: Technical Details

B.1. The Role of Preferences

The mean-variance preferences in the main text follow from investors with CARA preferences for terminal wealth and a preference for early resolution of uncertainty. Written as a function of terminal wealth, $U = E\left[-\log\left(E\left[\exp\left(-\rho W\right)|\hat{\mu},\mu\right]\right)|\mu\right]$. This form of preferences is not needed for the substitutability force that makes investors want to learn about different risks. It is required for the specialization result, where investors learn about one risk factor, in order to reduce their portfolio risk more efficiently.

These preferences make sense when considering information choice because investors with standard expected utility are not averse to the risk they face at the time when they form their portfolios. 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. Peress (2009) uses these same preferences for the same reason.

Because standard expected utility delivers no preference for early resolution of uncertainty, we use a utility function that makes early resolution desirable, similar to Epstein-Zin preferences. This formulation keeps the exponential structure of preferences that makes the problem tractable. It also has the advantage that learned information and prior information have the same effect on expected utility.

If we instead use standard expected utility, the learning choice is indeterminate (see Van Nieuwerburgh and Veldkamp (2008), Proposition 1). 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 an arbitrarily small preference for early resolution of uncertainty.

Using these preferences is not in any way assuming the result. The fixed-portfolio model in Section II 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. Our investors are asking, "What variance of beliefs would I most like to have in period 2, when I decide how to invest?"

An alternative interpretation of our solution is that we are solving a standard expected utility model, resulting in multiple equilibria. We select the equilibrium that gives the investor the highest period 2 utility. Using anticipatory utility, as in Brunnermeier and Parker (2006), would result in the same equilibrium selection. Another way to break indifference is to use constant *relative* risk aversion (CRRA) preferences. The problem with CRRA is that the model is no longer analytically tractable. The numerical solution of the CRRA partial equilibrium problem produces specialization for every set of parameters we tried.

B.2. Relaxing the Risk Factor Assumption

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.

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 then becomes

$$\max_{\hat{\Lambda}} \frac{1}{2} Tr\left(\widehat{\Gamma}\hat{\Lambda}^{-1}\widehat{\Gamma}'\Sigma_p - I\right) + \frac{1}{2}E[f-pr]'\widehat{\Gamma}\hat{\Lambda}^{-1}\widehat{\Gamma}'E[f-pr]. \tag{IA.5}$$

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

Strategic substitutability persists with arbitrary risk factors: Next, we show that investors still prefer to learn about different risks. 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 formulas, (A1), (A2), and (A3), will have the same forms as in Appendix B of the paper, 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 (IA.5) in two places: through Σ_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 ith 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 Λ_{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 in the exploitable pricing error. The last step links this to the average investor's uncertainty: $\partial\Lambda_{pjj}/\partial\Lambda^a_{\eta j} > 0$ because equation (A5) tells us that $\Lambda_{pjj} = (\Lambda^a_\eta)^2$, times a positive constant, and Λ^a_η is always positive. Thus, the first component of the value of learning about $\widehat{\Gamma}_i$ is always weakly increasing in aggregate uncertainty $\Lambda^a_{\eta j}$. 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 (A1) reveals the expected return on aggregate risk j: $\tilde{\Gamma}'_j E[f - pr] = \rho[\frac{1}{\rho^2\sigma_x^2}(\Lambda^a_{\eta j})^{-2} + (\Lambda^a_{\eta j})^{-1}]^{-1}\tilde{\Gamma}'_j \bar{x}$. This converges to zero as the average investor's precision of beliefs $(\Lambda^a_{\eta j})^{-1}$ 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 $(\Lambda^a_{\eta j})^{-1}$ increases.

In other words, no matter what risk factor an investor learns about, when others learn about correlated risks, it reduces the expected returns and exploitable pricing errors on that risk factors and makes it less desirable for to learn about. Strategic substitutability does not depend on how that risk factor is constructed.

B.3. Introducing Endogenous Capacity Choice

The choice of how much capacity to acquire would not change the decision of how to allocate that capacity, as long as cost was any increasing function of the reduction in generalized variance. This is simply a duality result. For any choice of capacity cost, there is a capacity endowment κ that produces identical outcomes. But once an amount of capacity is chosen, it will be allocated according to the results in this paper. Endogenizing capacity can produce other insights. Peress (2004) and Turmuhambetova (2005) show that the quantity of information chosen can explain realistic features of the relationship between income and investment patterns. But that problem can be neatly decomposed from our capacity allocation problem.

C. Connecting the Model to the Empirical Evidence

This section looks at how the model connects precisely to the empirical literature on portfolio outperformance. Outperformance is measured as either excess return or excess risk-adjusted return. We show that depending on how that risk adjustment is done, the results may or may not be a valid test of our model.

By outperformance, we mean the profits on a portfolio of assets. Informed investors do not earn higher returns on any given asset. Their outperformance comes from the covariance of the portfolio holdings with the asset returns $cov(q, f - pr)$. If this covariance is high, it means that the investor is buying assets that are likely to have high returns and selling ones that are likely to have low returns. In other words, outperformance arises because good information allows investors to pick the right assets to buy. In Hau (2001), Dvorak (2007), and Choe, Kho, and Stulz (2005), out-performance is based on such portfolio returns, connecting cleanly with our theory.

C.1. High Capacity Investors Earn High Returns

The following shows formally that higher capacity investors earn higher expected returns on their portfolios.

As in Admati (1985), we define the return on asset i as $(f_i - p_i r)$. Similarly, we define the expected portfolio return to be $q'(f - pr)$. The expected portfolio q is the only term in the expected return that is affected by an individual's K: one individual's capacity does not affect the aggregate variable A that determines $f - pr$, or the given exogenous risks Γ or prior beliefs Λ.

Furthermore, within the optimal portfolio, only the holdings of assets that load on the risk factor the investor learns about are affected by the investors capacity. Let i be the risk factor that the investor learns about: $i = \arg\max_{\ell \in \{1,\cdots,N\}} \mathcal{L}^j_\ell$. To see why this is true, we pre-multiply the optimal portfolio (equation (4) above) by the risk factor weights for risk factor i: $E[\Gamma'_i q] = \hat{\Lambda}_i^{-1}\hat{\Lambda}_i^a \Gamma'_i \bar{x}$. By Proposition 2, the investor always devotes his capacity to one risk factor. Therefore, capacity allows the investor to increase his posterior precision for the risk factor he learns about ($\frac{\partial \hat{\Lambda}_i^{-1}}{\partial K} > 0$), but does not affect the posterior variance of any other risk factor j. Thus, it does not affect $E[\Gamma'_j q]$ for any $j \neq i$.

It follows that capacity increases the investor's expected portfolio return if and only if the return on the component of the portfolio that loads on risk factor i increases. The expected portfolio return on a risk factor i is $E[(\Gamma'_i q)'\Gamma'_i(f - pr)]$. By the definition of covariance, this expected return can be expressed as the expected portfolio weight times the expected return of the factor, plus their covariance:

$$E[(\Gamma'_i q)'\Gamma'_i(f - pr)] = E[\Gamma'_i q]' E[\Gamma'_i(f - pr)] + \Gamma'_i cov(q, f - pr)\Gamma_i. \qquad \text{(IA.6)}$$

Using $E[\Gamma'_i q] = -\rho^{-1}\hat{\Lambda}_i^{-1}\hat{\Lambda}_i^a \Gamma'_i A$ and $f - pr = -A - Cx$, it can be shown that $E[\Gamma'_i(f - pr)] = -\Gamma'_i A$ and $cov(q, f - pr) = \frac{1}{\rho}(\widehat{\Sigma}^{-1}\Sigma - I)$. Canceling out orthogonal eigenvectors, we can rewrite

$$E[(\Gamma'_i q)'\Gamma'_i(f - pr)] = \frac{1}{\rho}\left(\hat{\Lambda}_i^{-1}(\Gamma'_i A)^2 + \hat{\Lambda}_i^{-1}\Lambda_i - 1\right), \qquad \text{(IA.7)}$$

Recall that capacity allows the investor to increase his posterior precision for risk i: $\frac{\partial \hat{\Lambda}_i^{-1}}{\partial K} > 0$. Since the expected return is increasing in this posterior precision, capacity increases the expected portfolio return.

C.2. Portfolio Concentration and Outperformance

Another popular way of empirically testing the asymmetric information hypothesis for portfolio underdiversification is to determine whether concentrated portfolios outperform diversified ones.

First, we show that investors with more underdiversified portfolios have higher capacity. Then, we invoke the argument from the previous section that links higher capacity to outperformance. Proposition 2 reveals that if investor j has higher capacity K, he uses it to further reduce the variance of $\hat{\Lambda}_i$ where $i = \arg\max_\ell \mathcal{L}^j_\ell$. Pre-multiplying both sides of equation (4) above by the risk-factor weighting matrix Γ reveals the link between this posterior variance and the optimal portfolio:

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

For all risk factors that the investor does not learn about, his portfolio holdings are unaffected by changes in his own capacity.

We define a more underdiversified portfolio to be one with a larger $|E[q] - \bar{x}|$. Pre-multiplying this distance by Γ_i and substituting in the formula for $\Gamma_i' E[q]$ yields $|(\hat{\Lambda}_i^{-1}\hat{\Lambda}_i^a - 1)\Gamma_i \bar{x}|$. Proposition 3 shows that in equilibrium, an investor is never less informed that the average investor in the risk factor he learns about. Therefore, $\hat{\Lambda}_i^{-1}\hat{\Lambda}_i^a > 1$. A higher K makes $\hat{\Lambda}_i^{-1}\hat{\Lambda}_i^a$ larger, which monotonically increases underdiversification. Thus, more capacity and more underdiversification are linked one-for-one in the model.

C.3. Do High Portfolio Returns Indicate Outperformance?

Ivkovic, Sialm, and Weisbenner (2007) and Massa and Simonov (2006) document that although more concentrated portfolios earn higher returns than diversified ones, these portfolio payoffs also have higher standard deviations. This leads other researchers (Guiso and Jappelli (2006)) to conclude that the returns come from excess risk taking. The model can shed light on whether high returns are indicative of information-based outperformance or risk-taking. All depends on how portfolio risk is measured. Risk is a standard deviation, which is a moment based on an expectation of a demeaned, squared random variable. The expectation is conditional on an information set. But this information set is not observed by the econometrician. According to the model, the traditional measure of portfolio risk, the unconditional standard deviation, is plagued by measurement error. Fortunately, the model also offers a solution to this measurement problem.

Risk is the conditional, not the unconditional standard deviation: Suppose there is a series of asset returns that follows the sequence $\{0.01, 0.10, 0.01, 0.10, 0.01 \ldots\}$. This sequence has a high standard deviation. Its unconditional Sharpe ratio is less than one. Now suppose that an investor knows with 100% certainty that this asset pays off 0.01 in odd periods and 0.10 in even periods. This is no longer a risky asset because its payoffs are certain. Conditional on that information, the standard deviation is zero. The conditional Sharpe ratio is infinite. An econometrician who did not know what the agent knows would compute historical moments of this asset payoff sequence and would deem a portfolio comprised of this asset a very risky portfolio.

Mapping this discussion back into the model, unconditional risk is the square root of the prior variance, $\Sigma^{1/2}$. Think of prior information as all past payoff realizations. When researchers measure outperformance, they are conditioning on this prior information. When investors can acquire private information, their expectations change and so does their conditional risk. Conditional risk is the square root of the posterior variance, $\widehat{\Sigma}^{1/2}$. This conditional risk is the same as the unconditional risk for the risks an investor did not learn about and is strictly less for risks the investor did learn about. The equation that links these two is

$$\hat{\Sigma}^j \equiv V[f|\mu^j, \eta^j, p] = \left((\Sigma^j)^{-1} + (\Sigma_\eta^j)^{-1} + \Sigma_p^{-1}\right)^{-1}. \tag{IA.8}$$

In other words, the relevant measure of risk for the investor depends on his information set, which the empirical researcher does not observe. In the language of our model, investor j's optimal portfolio $q^j = \frac{1}{\rho}(\widehat{\Sigma}^j)^{-1}(\hat{\mu}^j - pr)$ is conditioned on the information the investor knows $\hat{\mu}, \widehat{\Sigma}$. This portfolio does not maximize the investor's utility based on priors μ and Σ.

More specifically, a test that looks at unconditional Sharpe ratios to determine if underdiversified portfolios reflect superior information should always reject the information hypothesis. Diversified portfolios always have higher unconditional Sharpe ratios. The same is true in our model. But the efficient frontier that the empiricist constructs is not the same as the investors' frontier. Informed investors can achieve outcomes that are unachievable without their information.

Theory helps to distinguish conditional from unconditional: The theory offers a way of estimating conditional standard deviations. We need to know two things about an investor: How much capacity does he have and what risk does he assign the highest learning index to? To estimate an individual's information capacity, we can use their excess portfolio return, as in Section IV.B. To estimate the learning index, we need to estimate the average investor's learning index, as in Section IV.D of the paper. Then, we need to adjust that index for individual characteristics, such as their location, profession, or what school they studied at. Using this information, we can construct an individual investor's information using the following relationships: $\widehat{\Sigma} = \Gamma\hat{\Lambda}\Gamma'$ and $\hat{\Lambda}_i = \Lambda_i/K$ for the risk i with the highest learning index and $\hat{\Lambda}_j = \Lambda_j$ for all other risks j. The construction of Γ and Λ is described in Section IV.D of the paper. Then, we can ask whether the investor's portfolio maximizes the Sharpe ratio, conditional on this information set.

References

Admati, Anat, 1985, A noisy rational expectations equilibrium for multi-asset securities markets, *Econometrica* 53, 629–657.

Brunnermeier Markus and Jonathan Parker, 2006, Optimal expectations, *American Economic Review* 95, 1092–1118.

Choe, Hyuk, Bong-Chan Kho, and Rene M. Stulz, 2005, Do domestic investors have an edge? The trading experience of foreign investors in Korea, *Review of Financial Studies* 18, 795–829.

Dvorak, Tomas, 2007, Do domestic investors have an information advantage? Evidence from Indonesia, *Journal of Finance* 60, 817–839.

Guiso, Luigi, and Tullio Jappelli, 2006, Information acquisition and portfolio performance, CEPR working paper #5901.

Hau, Harald, 2001, Location matters: An examination of trading profits, *Journal of Finance* 56, 1959–1983.

Ivkovic, Zoran, Clemens Sialm, and Scott Weisbenner, 2007, Portfolio concentration and the performance of individual investors, *Journal of Financial and Quantitative Analysis* 43, 613-656.

Massa, Massimo, and Andrei Simonov, 2006, Hedging, familiarity and portfolio choice, *Review of Financial Studies* 19, 633–685.

Peress, Joel, 2004, Wealth, information acquisition and portfolio choice, *Review of Financial Studies* 17, 879–914.

Peress, Joel, 2009, The tradeoff between risk sharing and information production in financial markets, *Journal of Economic Theory* forthcoming.

Turmuhambetova, Gauhar, 2005, A simple portfolio problem with endogenous information, Working paper, University of Chicago.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2008, Information acquisition and under-diversification, Working paper, New York University.