Abstract

Observed portfolios are concentrated in asset classes which comove strongly with the non-financial income of investors. As an explanation, I propose a framework of endogenously generated information asymmetry, where rational agents optimally choose to focus their limited attention on risk factors that drive both their non-financial income and some of the risky asset payoffs. In turn, the agents concentrate their portfolios in assets driven by those factors. The paper also shows that exogenous information asymmetry cannot deliver the result and that a theory of endogenous information asymmetry is key for an information based explanation of portfolio bias.

JEL Codes: F3, G11, G15, D8, D83
1 Introduction

Empirical evidence indicates that investors fail to hold diversified portfolios and appear to leave themselves exposed to substantial amounts of, otherwise avoidable, idiosyncratic risk. Furthermore, portfolios are concentrated in assets that are strongly positively correlated with investors' non-financial income, and thus appear to provide little hedging benefits. In contrast to this, traditional economic theory suggests that rational agents will hold a large variety of assets exposed to different sources of risk, and in particular favor assets which are not positively correlated with their non-financial income. These systematic differences between observed and model-predicted portfolios is usually referred to as “portfolio bias”, and is a long-standing puzzle in both financial and international economics.

The empirical literature has documented a number of portfolio biases, with the “equity home bias” being perhaps the most popular example. The term refers to the observation that aggregate national portfolios are strongly biased towards holding domestic equity assets and generally exhibit a low propensity for investing in foreign equity assets (see among others French and Poterba (1991), Tesar and Werner (1998), and Ahearne et al. (2004)). A number of other portfolio biases have also been identified at the individual investor level. Huberman and Sengmueller (2004), Poterba (2003), Benartzi (2001), find that investors overinvest in the stock of their own employer or in the stock of other companies in the same industry. Ivković and Weisbenner (2005) and Huberman (2001) find that investors exhibit a “local bias” – the tendency to overinvest in assets that are geographically close to the investor's place of residence. Massa and Simonov (2006) show that investors tilt their portfolios away from the market portfolio, and towards assets with a significantly higher correlation with their non-financial income. Goetzmann and Kumar (2008) show that the stocks individual investors hold are highly correlated with each other, which suggests that they are driven by the same risk factors.

This paper proposes a framework of endogenously generated information asymmetry that can help resolve the puzzle. The information asymmetry arises because the agents choose to be better informed (i.e. have a higher precision in their rational beliefs) about some risk factors as compared to others. The idea that information asymmetry can generate concentrated portfolios is quite intuitive, and goes back to Merton (1987) who imposed the information structure exogenously.
Recent work by Van Nieuwerburgh and Veldkamp (2009, 2010) builds on that result by allowing the agents to make optimal information acquisition choices. In their framework the information choices of agents amplify any pre-existing information differences but cannot generate information asymmetry, and hence portfolio bias, if agents have symmetric prior information. I extend the literature by developing a model in which rational agents, who possess the same prior information about all assets, choose learning strategies that generate ex-post information asymmetry. Moreover, I also show that some of the previous literature results appear to be knife edge cases and that a model based solely on exogenously given prior information differences cannot, in general, generate portfolio concentration. I find that a framework where endogenous economic forces compel agents to value information differently, such as the one presented in this paper, is necessary for obtaining information asymmetry that is robust to model perturbations.

The key ingredients to the model are non-diversifiable non-financial income (incomplete-insurance) and information processing constraints. Households can observe signals that provide noisy information about the future states of the world, but unlike the traditional signal extraction literature the structure of the signals is not imposed exogenously but is determined endogenously. The households choose the precision of their signals, subject to the constraint that the signals can only contain a finite amount of information. The information acquisition framework is a version of the Rational Inattention framework developed by Sims (2003). The information constraint compels agents to focus their attention on information that is highly pertinent to their particular situation. It turns out that the agents are most sensitive to factors that affect both their non-financial and their financial incomes and hence focus most of their attention on them. In turn, this information acquisition behavior generates information asymmetry where agents end up possessing information of different quality about the different factors that can affect asset payoffs. As a consequence, the optimal portfolio choice of the agents leads them to concentrate their holdings in assets closely related with their non-financial income. It is important to note that the portfolios are the result of optimal behavior, even though they may look inefficient when viewed through the lenses of a standard, symmetric information model.

An important implication of the model is that portfolio concentration has
a non-monotonic relationship with an agent’s capacity to process information. Agents with a low capacity find it optimal to specialize in acquiring information about only one asset and thus allocate any extra information capacity to the same asset, while agents with a high capacity choose to spread their efforts across a variety of different assets. Thus, the resulting information asymmetry and hence the degree of portfolio concentration is increasing in the capacity for information acquisition when the capacity is low, and decreasing when it is high. On the other hand, existing models that allow for optimal information acquisition but are based on exogenous prior information differences imply that the agents always choose to specialize their information acquisition in a single asset. In such models, the relationship between portfolio concentration and the capacity for acquiring information is unambiguously positive. This and other differences in implications allows me to empirically differentiate between the two frameworks. In Section 6 I show that the model of this paper is consistent with the data, while the implications of models based on exogenous prior differences are rejected. In this sense, a model of endogenous information asymmetry is key for an information based explanation of portfolio concentration, as mechanisms that generate information asymmetry through prior information differences lead to counterfactual implications.

In addition, the model has the implications that portfolio concentration is generated by information asymmetry and that the information asymmetry is tilted towards assets associated with the investors’ non-financial income. These features of the model are also supported by the empirical evidence. For example, Bae et al. (2008) find that domestic analysts make better predictions than foreign analysts and Malloy (2005) finds that when considering a single country, local analysts outperform out-of-town analysts. On the other hand, as was already discussed, a number of papers show that investors concentrate their portfolios in the stock of local companies and/or the companies that employ them. Moreover, Ivković and Weisbenner (2005), Massa and Simonov (2006) and Ivkovic et al. (2008) document that investors earn excess returns on their local holdings, suggesting that they do possess superior information over the assets they hold.

The paper is organized as follows. Section 2 discusses the related literature. Section 3 presents the details of the model and Section 4 discusses its key ingredients. Section 5 presents the solutions to the model and derives the main analytical results.
Lastly, Section 6 presents the empirical evidence and Section 7 concludes.

2 Information Frictions, Nontradeable Labor Income and Portfolio Choice

This paper is related to three different strands of literature: the literature on portfolio choice under exogenous information asymmetry, the literature on portfolio choice and endogenous information acquisition, and the open economy macroeconomics literature relating home bias and labor income.

Assuming an extreme form of information asymmetry exogenously, Merton (1987) shows that when investors have heterogeneous information sets individual portfolios can differ significantly from the benchmark representative agent portfolio. A number of subsequent papers (Gehrig (1993), Brennan and Cao (1997), Brennan et al. (2005), Hatchondo (2008), and others) expand on this insight and apply it specifically to the home bias equity puzzle by developing two-country models, where home and foreign agents receive signals about the future payoffs of assets, but home agents are assumed to receive more precise signals about the home asset. The major drawback of this approach is summarized by Pástor (2000), who shows that for sufficient home bias to exist, the home agents must possess very strong prior information advantages. It is not clear why agents who start with a large information advantage over domestic asset would not spend resources to inform themselves about the highly uncertain (in relative terms) foreign asset. This paper provides an answer to this criticism by developing a framework where information asymmetry is generated by the endogenous learning choice of investors.

Van Nieuwerburgh and Veldkamp (2009, 2010), Mondria (2010) and others, extend the previous literature on exogenous information asymmetry by allowing for endogenous information acquisition choice. Models in this mold exhibit increasing returns to information acquisition, and thus the optimal attention allocation strategy is the corner solution where agents only learn about one of the assets. Portfolio concentration is then generated by endowing the agents with an arbitrarily small information advantage over one of the assets, which ensures that the optimal solution is to learn only about that asset. The results crucially depend on both the
increasing returns to information acquisition and the assumption of exogenous prior information asymmetry. However, as I show in Section 6, the increasing returns to information acquisition lead to counter-factual implications. This paper adds to the previous work by introducing a mechanism that does not rely on increasing returns to information or exogenous prior differences, and is thus able to endogenously generate information asymmetry in a way that agrees with the data.

Mondria and Wu (2011) extend Mondria (2010) by considering imperfectly integrated international financial markets. In their framework, information asymmetry can also arise endogenously, but is driven by the fact that agents face transactions costs only when they trade foreign assets, but not when they trade home assets. The results are dependent on the asymmetric transaction costs associated with home and foreign assets, which could be justified in an international setting but are harder to motivate in a closed economy setting. Furthermore, Mondria and Wu (2011) use the previous literature’s information acquisition framework, which always exhibits increasing returns to information acquisition, and thus has counter-factual implications.

Lastly, the paper is also related to the open macroeconomics literature on the home bias, and specifically the strand of literature that considers the importance of labor income in the determination of international portfolios. Coeurdacier and Gourinchas (2011) and Heathcote and Perri (2007) develop two distinct frameworks where the joint determination of the equilibrium real exchange rate, labor income and asset returns generates a positive labor income-hedging demand for the home equity asset. This paper shares the literature’s key insight that labor income plays an important role in the formation of home biased portfolios, but the mechanisms are fundamentally different. In this work, labor income does not provide a positive hedging demand, but rather is the reason that the agents decide to bias their information acquisition strategy towards the home asset. The two frameworks are complimentary and would amplify each other’s results, if put together. An important difference, however, is that the current model can simultaneously explain the portfolio concentration in both international and domestic portfolios, while the hedging motive identified by the macro literature only operates in an international environment with volatile exchange rates.
3 Model Framework

This paper models the information acquisition decision and asset demands of rational agents who face information processing constraints and take prices as given. For ease of exposition, I will present the model in an international setting, assuming that agents choose between “home” and “foreign” assets. This allows for the unambiguous terminology “home” and “foreign” but all results carry over to a closed, multi-sector (region) economy where an agent has the option of investing in different sectors (regions) of the economy.

There is a continuum of agents of mass 1 that live in the home country, earn their non-financial income there, and can trade a riskless bond and “home” and “foreign” risky securities. The agents live for 3 periods – they make information acquisition choices in period 0, observe informative but noisy signals and then choose their portfolio allocations in period 1 and in period 2 shocks are realized and the agents consume all of their wealth. Agents maximize utility over period 2 consumption, and do not value leisure. Therefore agents supply their whole labor endowment, normalized to 1, at the wage rate they face. Thus each agent faces the following period 2 budget constraint,

\[ c_2 = \delta w + x_h y_h + x_f y_f + b R, \]

where \( w \) is the wage rate, \( y_h \) is the payoff of the home asset, \( y_f \) is the payoff of the foreign asset, \( R \) is the gross return on the riskless bond, and \( x_h \) and \( x_f \) are the quantities of the home and foreign asset the agent chose in period 1. The first term in the above equation, \( \delta w \), is the non-financial (wage) income of the individual, and the rest is his financial based income. The parameter \( \delta \) controls the relative weight of labor income in the agent’s total income. When \( \delta = 0 \) the agent has no labor income and period 2 consumption relies entirely on financial income, and the importance of labor income increases as \( \delta \) grows. The existence of labor income is

---

1In this paper I model the non-financial income as labor income, but in actuality this is the income stream coming from all other sources, but the publicly traded financial investments of an agent. Hence, this also includes the profits/dividends received from any privately-owned companies, private investments and etc.

2Given the static nature of the model, the agents do not make a decision on changing the country (industry) they work in (for) or the amount of time spent on working. An infinite horizon extension of the model which allows for these considerations is left to future work.
central to the mechanism of the model, and I will inspect how the relative size of non-financial to financial income affects the results by varying $\delta$.

The focus of the paper is on the information acquisition behavior of the households (investors) and is silent on the equilibrium determination of wage and rent rates, hence I model the wage and the payoffs to the risky assets as exogenous stochastic processes. I assume that the 3 payoff processes are given by:

$$w = z_h + \varepsilon_w$$
$$y_h = z_h + \varepsilon_h$$
$$y_f = z_f + \varepsilon_f$$

where all variables are normally distributed with $[z_h, z_j]' \sim N(\mu, \Sigma)$, $\mu$ is a column vector with $\mu_1 = \mu_2$, and all epsilons i.i.d. $N(0, \sigma_j^2)$, $j \in \{w, h, f\}$. Each payoff is driven by two stochastic factors, and the factor $z_h$ drives both labor income, $w$, and the payoff to home equity $y_h$. This captures the idea that labor income and home equity are both affected by some of the same forces.\(^3\) However, the two are not perfectly correlated as the factors $\varepsilon_w$ and $\varepsilon_h$ are independent from each other (this ensures labor income is not traded and hence is not perfectly diversifiable). For simplicity, the framework also assumes $\Sigma$ is diagonal and hence there is no relationship between home and foreign payoffs. It is straightforward to introduce a “global” factor which would drive all three processes and thus introduce correlation across countries (or sectors), however I found this to have no qualitative effect on the results. For ease exposition, therefore, I present the model without such a “global” factor.

In period 1, an agent chooses her portfolio allocations subject to her initial budget constraint,

$$A = p_h x_h + p_f x_f + b,$$

where $A$ is the initial wealth of agent $i$, $p_h$ is the price of the home asset and $p_f$ is the price of the foreign asset. The paper analyzes the agents’ optimal information and portfolio choices in a partial equilibrium setup and keeps the prices fixed. It is straightforward to obtain the equilibrium prices following the results in Admati\(^8\)

\(^3\)One can think of $z_h$ as home country/sector TFP and of $z_f$ as the foreign TFP.
(1985), however I found that this has no effect on the optimal behavior of the agents, which is the main focus of the paper. Hence, for ease of exposition, I abstract from the equilibrium determination of the asset prices.

Prior to making the portfolio allocation choice the agents receive informative, but noisy, signals about \{z_h, z_f\} but not about any of the \(\varepsilon_j\). This captures the idea that payoffs are subject to two types of uncertainty – uncertainty that an agent can reduce by acquiring and processing information available today (newspapers, Federal Reserve announcements, Analyst reports, etc.) and uncertainty that the agents cannot reduce with any of today’s information. I will continuously refer to the first type as “learnable” uncertainty and to the other one as “unlearnable”. Without such a dual structure of uncertainty the model would imply that an agent with sufficiently large information processing capacity could forecast the future arbitrarily well. Given how hard it is to forecast macroeconomic and, especially, financial series, this seems unrealistic. Later sections will study how the existence of unlearnable uncertainty affects the results in full detail.

In particular, the agents receive unbiased signals of the form:

\[ \eta_h = z_h + u_h \]
\[ \eta_f = z_f + u_f \]

where \(u_h\) and \(u_f\) are independent of each other and \(u_h \sim N(0, \sigma^2_{u_h}), u_f \sim N(0, \sigma^2_{u_f})\).

Moreover, the agents are allowed to choose \(\sigma^2_{u_h}\) and \(\sigma^2_{u_f}\), which control the amount of noise in the signals \{\(\eta_h, \eta_f\}\}. This choice is made subject to the constraint that the total amount of information carried in the chosen signals is limited by the following entropy reduction constraint:

\[ H(z_h, z_f) - H(z_h, z_f | \eta_h, \eta_f) \leq \kappa, \]

where \(H(X)\) is the entropy of random variable \(X\) and \(H(X|Y)\) is the entropy of

---

4In a dynamic setting, however, the equilibrium determination of the price will play a key role. In that setting the agents would want to forecast future prices, which are determined simultaneously with the information choices of the market participants, and thus there are interesting feedback effects between agents’ decisions and the market price. The dynamic case is thus different, but is outside of the scope of this paper and left to future work.
X conditional on knowing $Y$. Entropy is the standard measure of uncertainty in information theory, and consequently reductions in entropy measure information flow. The expression $H(z_h, z_f) - H(z_h, z_f|\eta_h, \eta_f)$, measures the information about $\{z_h, z_f\}$ carried by $\{\eta_h, \eta_f\}$ and the above constraint states that agents can only choose signals that carry no more than a total of $\kappa$ bits of information.

The independence of $u_h$ and $u_f$ implies that agents can only reduce the posterior variance of $z_h$ and $z_f$ but cannot choose a correlation structure for their posterior beliefs – the correlation structure is taken as given from their priors. Intuitively, this amounts to assuming that learning about $z_h$ and $z_f$ are independent and unrelated activities that the agents carry out separately from one another, which appears to be a natural assumption given that the factors $z_h$ and $z_f$ themselves are assumed to be independent. After observing the described signals, the agents use Bayesian updating with the correct prior to form their posterior beliefs. As opposed to other papers in the literature, I assume the agents have identical priors over both the home and foreign factors. In this sense, there is no exogenously imposed information asymmetry and the rest of the paper focuses on showing that even without prior exogenous information advantages, the home agents’ utility maximizing decisions lead to ex-post information asymmetry.

Lastly, I assume that the agent’s preferences over period 2 consumption are summarized by the following utility function:

$$E_0(-\ln(-E_1(-\exp(-\gamma c_2))))$$

This is the “mean-variance” utility function used by the previous literature on information acquisition and portfolio choice (e.g. Van Nieuwerburgh and Veldkamp (2009, 2010)). It is a CARA utility function with an added desire for early resolution of uncertainty. This follows from results in Kreps and Porteus (1978) who show that if $u(c)$ is a standard Von Neumann Morgenstern utility function and $T(X)$

---

5 Entropy is defined as $H(X) = -E(\ln(f(x)))$, where $f(x)$ is the probability density function of $X$.

6 I could relax the assumption that signals are independent by allowing the agents to observe a linear combination of the factors $z_h$ and $z_f$ as in Mondria (2010), but I would lose the analytical tractability of the model. On the other hand, numerical solutions show that the main conclusions of the model and the nature of the mechanism examined remain unaffected. Thus, I maintain the assumption of independent signals.
is a convex operator, then $U = E_1(T(E_2(u(c))))$ is a version of the original utility function which exhibits a desire for early resolution of uncertainty, while preserving the original utility function’s risk aversion characteristics. The natural log is the typical choice for the convex operator $T(X)$ in the literature as it allows for tractable analytical solutions.

4 Key Ingredients

Before presenting the main results, it is useful to take a closer look at the key ingredients of the model and obtain some intuition about the mechanism which generates the information asymmetry. The basic intuition is that the agent’s period 2 consumption is more heavily dependent on $z_h$ than on $z_f$. Thus, home information is more pertinent to the home agent than foreign information, and consequently he focus his limited attention on it.

To see this in more detail, let $\bar{x}_h = x_h + \delta$ and rewrite the second period budget constraint as:

$$c_2 = \delta(\varepsilon_w - \varepsilon_h) + \bar{x}_h y_h + x_f y_f + b R$$

Now define $\bar{A} = A + \delta p_h$ and rewrite the initial wealth constraint in terms of the new variable $\bar{x}_h$:

$$\bar{A} = p_h \bar{x}_h + p_f x_f + b$$

This fully re-parameterizes the model in terms of the new choice variable $\bar{x}_h$ and allows us to acquire intuition about the mechanism more easily.

From the expression above, it is clear that $Cov(y_h, \delta(\varepsilon_w - \varepsilon_h)) = Cov(z_h + \varepsilon_h, \delta(\varepsilon_w - \varepsilon_h)) < 0$, and hence the home asset correlates negatively with non-financial income. Thus, the agent has two incentives to buy the home asset: 1) it offers high expected returns (risk premium) and 2) it can help hedge non-financial income risk. On the other hand, the foreign asset has zero correlation with the agent’s labor income, and thus the agents only buy it because of the high expected return. Because of the negative correlation between the home asset and the agent’s labor income, standard portfolio choice theory tells us that $\bar{x}_h > x_f$. Taking $\bar{x}_h$ and $x_f$
as given, it can be shown that the variance of consumption in period 2 is:

$$\text{Var}(c_2) = \delta^2 (\sigma_{\epsilon_w}^2 + \sigma_{\epsilon_h}^2) + \bar{x}_h^2 \text{Var}(y_h) + x_f^2 \text{Var}(y_f) - 2\delta \bar{x}_h \sigma_{\epsilon_h}^2$$

The volatility of $c_2$ is affected by the variance of $y_h$ more so than the variance of $y_f$ - second period consumption has a higher exposure to risk related to $y_h$ (and hence $z_h$) than to risk related to $y_f$ (and hence $z_f$).

We can conclude that there are two reasons for the agent to value the information contained in the signals $\eta_h$ and $\eta_f$. First, high precision signals would help the agent choose a portfolio allocation that maximizes risk-adjusted returns. Second, the signals will resolve some of the uncertainty about period 2 consumption, which is valued by the agent because the utility function exhibits a desire for early resolution of uncertainty. In regards to portfolio selection, the precision of $\eta_h$ is just as valuable to the agent as the precision on $\eta_f$. The CARA investor being studied here chooses portfolio allocations that simply optimize the mean-variance properties of the available assets. The two signals are equally useful in determining the mean-variance frontier and thus have symmetric effects on his expected utility. However, $\eta_h$ is more effective in reducing the posterior variance of $c_2$ than $\eta_f$, because as we showed above, $c_2$ has a higher exposure to the factor $z_h$ than to $z_f$. Thus, the optimal choice of information acquisition puts a higher precision on $\eta_h$ than on $\eta_f$, because a unit of information about $z_h$ eliminates more uncertainty about $c_2$ than a unit of information about $z_f$.

Lastly, note that the desire for early resolution of uncertainty is key in generating endogenous information asymmetry in this model but is not necessary in general. All that is needed is for the agent to value information about his future non-financial income. A tractable, though a bit mechanical, way to achieve this is through the introduction of a desire for early resolution of uncertainty. Another way is to use a model where knowledge about future non-financial income informs the agent’s actions today and is thus useful in today’s optimization. For example, this happens in a dynamic model where the agent is also making a savings decision, as savings depend on the expectation of future income. It also arises under standard power utility because the agent then has decreasing absolute risk aversion, and the expectation of future non-financial income factors in the agent’s appetite for risk and thus his optimal portfolio. These setups do not yield themselves to analytical
solutions, however, and I have opted for the setup of this paper precisely because it allows for the derivation of analytical results. Nevertheless, numerical solutions of the other setups yield the same qualitative results and I leave the detailed analysis of such extensions to future work.\footnote{Regular CARA utility also yields itself to analytical solutions. In that case, I find that investors value home and foreign information equally and there is no information asymmetry. This result arises because home and foreign information are equally useful in determining the mean-variance frontier, as discussed in the previous paragraph.}

5 Model Solution

5.1 Period 1: Portfolio Allocation

It is convenient to define the following notation for the posterior and prior variances of $z_h$ and $z_f$: $\text{Var}(z_h|\eta_h) = \sigma_h^2$, $\text{Var}(z_f|\eta_f) = \sigma_f^2$, and $\text{Var}(z_h) = \text{Var}(z_f) = \sigma_z^2$. Having observed signals $\{\eta_h, \eta_f\}$, the agents use Bayesian updating and their prior beliefs to obtain the posterior distribution $\hat{z}|\hat{\eta} \sim N(\hat{\mu}, \hat{\Sigma})$, where $\hat{\mu} = E(\hat{z}|\hat{\eta})$ and $\hat{\Sigma} = \begin{bmatrix} \sigma_h^2 & 0 \\ 0 & \sigma_f^2 \end{bmatrix}$. Taking the conditional distribution as given, the solution to an agent’s period 1 problem yields the standard CARA utility portfolio choice rules:

$$x^*_h = \frac{\hat{\mu}_h - p_h R}{\gamma(\sigma_h^2 + \sigma_z^2)} - \frac{\delta \sigma_h^2}{\sigma_h^2 + \sigma_z^2}$$

$$x^*_f = \frac{\hat{\mu}_f - p_f R}{\gamma(\sigma_f^2 + \sigma_z^2)}$$

The expressions show two things. First, an agent has an additional hedging motive to trade the home asset (the second term in the expression for $x_h$) that does not factor into the demand for the foreign asset. Secondly, one can already see how portfolio concentration can be obtained if $\sigma_h^2 \neq \sigma_f^2$. The optimal choice of $\sigma_h^2$ and $\sigma_f^2$ is the focus of the next section.

5.2 Period 0: Information Acquisition Choice

In period 0 the agent takes into account the form of his optimal portfolio choice allocations in period 1 and chooses the utility maximizing posterior variances for
the factors $z_h$ and $z_f$, such that the signals he receives carry no more than $\kappa$ bits of information.\textsuperscript{8} The main result of the paper is that the optimal information acquisition choice is always such that the agent acquires more information about $z_h$ than about $z_f$. This is formalized in Proposition 1 below.

**Proposition 1.** Let i) $\delta > 0$ and ii) $\sigma^2_{u_h} > 0$. Then the unique solution to the information acquisition problem, $\{\sigma^2_h, \sigma^2_f\}$, is such that $\sigma^2_h < \sigma^2_f$, i.e. a home agent always pays more attention to the home forecastable factor than to the foreign forecastable factor.

**Proof.** The proof is in the Appendix. \hfill \Box

Proposition 1 summarizes the main result of the paper. In the presence of non-diversifiable labor income, the optimal signal about the home factor has higher precision than the optimal signal about the foreign factor. The model endogenously generates the information asymmetry which is typically introduced exogenously by the previous literature.

The key for the result is the existence of a non-diversifiable, forecastable component of non-financial income. Condition i) states that the agent has any non-financial income at all ($\delta > 0$), and condition ii) ensures that the forecastable component is non-diversifiable given the available financial instruments. To see the second point, imagine that ii) did not hold. Then the agent will be able to trade the risk factor $z_h$ and since non-traded income is linear in $z_h$ he will be able to perfectly hedge all forecastable risk in it. The only remaining uncertainty in the non-financial income will then be unrelated to the forecastable factors $z_h$ and $z_f$ and will thus have no effect on the information acquisition decision of the agent.

It is important to highlight that while crucial, non-financial income does not do all the heavy lifting by itself. An amplification mechanism similar to the one first identified by Van Nieuwerburgh and Veldkamp (2010) is at work in the following way. The forecastable component of non-financial income induces the agent to shift some of his attention toward the home factor, which in turn translates to lower

\textsuperscript{8}Technically, the agent chooses the variances of the error terms in his signals, $\sigma^2_{u_h}$ and $\sigma^2_{u_f}$. However, there is a one-to-one relationship between the posterior variances of the $z$ factors and the variance of the error terms, $\sigma^2_h = (\frac{1}{\sigma^2_f} + \frac{1}{\sigma^2_{u_h}})^{-1}$ and $\sigma^2_f = (\frac{1}{\sigma^2_f} + \frac{1}{\sigma^2_{u_f}})^{-1}$, and it is more convenient to parameterize the problem in terms of $\sigma^2_h$ and $\sigma^2_f$. 

14
posterior variance of the home risky asset return as compared to the foreign risky asset. This compels the agent to rebalance his portfolio towards the home asset but this increases his dependence on the home factor even more. This provides him with an even greater incentive to shift attention toward the home factor, which leads to further portfolio rebalancing and so forth.

A direct implication of Proposition 1 is that portfolio allocations will diverge from what traditional theory would consider as a “well diversified” portfolio. Formally we have the following Corollary:

**Corollary 1.** Let \( \{x^B_h, x^B_f\} \) be the mean portfolio allocations of an economy of Bayesian agents who receive equally precise signals about \( \{z_h, z_f\} \) in period 1. Let \( \{x^O_h, x^O_f\} \) be the allocations under optimal attention allocation, where the signals are chosen so that they carry the same amount of total information as in the previously defined economy. Then, \( x^O_h > x^B_h \) and \( x^O_f < x^B_f \) – optimal information acquisition biases portfolios towards the home asset.

**Proof.** Follows directly from Proposition 1 and the portfolio choice expressions. 

This Corollary is not surprising since Proposition 1 tells us the model generates information asymmetry and the previous literature has already formalized the relationship between information asymmetry and portfolio concentration. This paper introduces a mechanism that generates the information asymmetry endogenously, but does not affect the portfolio choice in any other way.

### 5.3 The Size of The Endogenous Information Asymmetry

In this section I characterize the size and behavior of the generated information asymmetry over the whole parameter space and find conditions under which it achieves its maximum and minimum values.

A natural measure for information asymmetry is the difference in the amount of information learned about the home forecastable shock, \( z_h \), and the amount of information learned about the foreign forecastable shock, \( z_f \). Call this difference

\[ \text{The unit of measure for information is bits, as is standard in information theory.} \]
in information flows $\Lambda$:

$$
\Lambda = (H(z_h) - H(z_h|\eta_h)) - (H(z_f) - H(z_f|\eta_f))
= \frac{1}{2} (\ln(\sigma_f^2) - \ln(\sigma_h^2))
$$

By Proposition 1, $\Lambda > 0$, or in other words, the agents always process more information about home shocks than about foreign shocks. Further, $\frac{\sigma_f}{\sigma_h} = \exp(\Lambda)$, thus there is a one-to-one positive relationship between the variable $\Lambda$ and the ratio of posterior standard deviations for the two forecastable shocks. Thus, increasing $\Lambda$, the measure of information asymmetry, decreases $\sigma_h$ relative to $\sigma_f$.

Before we study the behavior of $\Lambda$ itself, I derive the following Proposition.

**Proposition 2.** The objective function $U$ is convex in home information when $\sigma_h^2 \geq \sigma_e^2$ and concave otherwise. Similarly it is convex in foreign information when $\sigma_f^2 \geq \sigma_e^2$ and concave otherwise.

**Proof.** The proof is in the Appendix. \(\square\)

This result tells us that an agent enjoys increasing returns to information acquisition whenever the size of the remaining learnable uncertainty (either $\sigma_h^2$ or $\sigma_f^2$) is greater than the size of unlearnable uncertainty $\sigma_e^2$. This suggests that there are regions of the parameter space where the model has corner solutions and the optimal strategy of the agents is to allocate all available attention to only one of the factors – this is the relevant case for models in the spirit of Van Nieuwerburgh and Veldkamp (2010). However, there will also be regions of the parameter space where the problem is concave and the agent will acquire positive amounts of information about both factors. In particular, specialization pays off whenever information acquisition can reduce the majority of the uncertainty about an asset. In cases where information acquisition can only reduce a minor part of the present uncertainty, the optimal strategy is to learn some about both factors. The model’s implications are rich in the sense that different conditions (e.g. amounts of uncertainty being faced,
size of the information capacity etc.) would imply different optimal information acquisition strategies. The next few propositions characterize how the optimal information acquisition strategy changes with the salient parameters of the model.

The following proposition characterizes how information asymmetry changes with \( \kappa \).

**Proposition 3.** Let \( \delta, \gamma \) and \( \sigma^2_{\epsilon h} = \sigma^2_{\epsilon f} = \sigma^2_{\epsilon} \) be given (symmetric assets), and satisfy the conditions of Proposition 1. Then,

(i) There exists a value \( \tilde{\kappa}(\gamma, \delta, \sigma^2_{\epsilon}) \) such that \( \Lambda \) is an increasing function of information processing capacity, \( \kappa \), whenever \( \kappa < \tilde{\kappa}(\gamma, \delta, \sigma^2_{\epsilon}) \) and a decreasing function of \( \kappa \) for \( \kappa \geq \tilde{\kappa}(\gamma, \delta, \sigma^2_{\epsilon}) \).

(ii) When \( \kappa \leq \tilde{\kappa}(\gamma, \delta, \sigma^2_{\epsilon}) \) agents choose to devote their whole attention to the home forecastable factor, hence \( \Lambda = \kappa(\gamma, \delta, \sigma^2_{\epsilon}) \), \( \sigma^2_h = \frac{\sigma^2_{\epsilon}}{\exp(2\kappa)} \) and \( \sigma^2_f = \sigma^2_{\epsilon} \).

The agent starts to allocate some of his attention capacity to the foreign factor once \( \kappa > \tilde{\kappa}(\gamma, \delta, \sigma^2_{\epsilon}) \).

(iii) As \( \kappa \to \infty \), \( \Lambda \to \frac{1}{2} \ln\left(\frac{A}{B}\right) > 0 \), where \( A \) and \( B \) are constants defined in the appendix.

**Proof.** The proof is in the Appendix.

Proposition 3 develops three results that are useful in determining the size of the generated information asymmetry. First, it tells us that information asymmetry is a non-monotonic function of the information processing ability of an agent, where it is increasing for low values of information capacity and decreasing for large ones. Second, it informs us that some agents (ones for which \( \kappa \leq \tilde{\kappa}(\gamma, \delta, \sigma^2_{\epsilon}) \)) feel so constrained in terms of information capacity, that they allocate all of it to the home forecastable factor and ignore information about the foreign factor. Third, information asymmetry converges to a positive limit as information capacity becomes infinite – thus information asymmetry exists even in the limit of unbounded information processing ability.

Figure 1 plots \( \Lambda \) and the ratio \( \frac{\sigma_f}{\sigma_h} \) on \( \kappa \). One can clearly see the two regions defined by Proposition 3. At first, the agent devotes all attention to the home asset and \( \Lambda \) is growing linearly with \( \kappa \). After a certain point he starts paying attention to
the foreign asset as well, information asymmetry unravels and eventually converges to a positive number. The second plot shows the ratio of the posterior standard deviations of \( z_f \) and \( z_h \). This ratio starts at 1 as \( \kappa \) is 0 and at first rises, then falls, and eventually converges to a number strictly greater than 1. Thus, the value of \( \kappa \) is key to determining the size of the generated information asymmetry. At the two extremes of scarce and abundant information capacity information asymmetry is smallest, and is highest for moderate values of \( \kappa \).

Figure 1: Information Asymmetry and Information Capacity

The result that \( \Lambda \) is first increasing and then decreasing with information capacity is intuitive. It is easy to imagine that people with little time and/or information resources would focus their attention on issues important to their particular situation, and ignore information about other things. For example, imagine that you had only five minutes to check the web page of your favorite news provider. Odds are you will go to the page for domestic/local news and simply scan the headlines for the major news, as this will give you the best overall picture
of the things that matter most to you that can be obtained in five minutes. Now imagine that you had twenty minutes to read the news. You will still go to the domestic/local news page first and scan the headlines, but then you will also most likely read through an entire article or two. This is an example of the increasing returns to information acquisition when information processing abilities are scarce. Reading just the title might give you an interesting piece of information like “The Economy is Close to a Recession”, but the body of the article would be even more useful (increasing returns). Finally, imagine that you have several hours to spend reading the news. Most likely, in such a large amount of time you will drive down the marginal benefit of domestic/local information to the point where you move on to the news page of a country that is of particular interest to you, for example possibly one in which you have invested some money. This is an example of the decreasing returns to information acquisition which occur when \( \kappa \) is sufficiently large.

The next Proposition derives results about how information asymmetry and the shape of the graph in Figure 1 change with risk aversion (\( \gamma \)), the size of labor income (\( \delta \)) and the total amount of unlearnable uncertainty (\( \sigma^2_\varepsilon \)).

**Proposition 4.** Let the conditions of Proposition 1 be satisfied. Then, \( \bar{\kappa}(\gamma, \delta, \sigma^2_\varepsilon) \) is an increasing function of \( \gamma, \delta \) and \( \alpha \), where \( \sigma^2_\varepsilon = \alpha \sigma^2_\zeta \). Moreover, \( \Lambda \) - the overall amount of information asymmetry - is an increasing function of \( \gamma \) and \( \delta \).

**Proof.** The proof is in the Appendix.

More risk averse agents and agents for whom labor income is a more important source of funds have stronger incentives to skew their attention towards the home factor. Moreover, in a world where most of the uncertainty agents face cannot be reduced through acquiring the information available today (i.e. low \( \alpha \)), the payoffs to information specialization are low.\(^{10}\) Hence, the attention tipping point \( \bar{\kappa} \) occurs at a smaller value of \( \kappa \), as compared to environments where the agents can use the information they acquire to reduce most of the uncertainty they face.

Figure 2 plots \( \Lambda \) and \( \frac{\sigma^2_\varepsilon}{\sigma_\kappa} \) on \( \kappa \) for three different values of the risk aversion coefficient \( \gamma \). The graph illustrates the result of Proposition 4 that information

---

\(^{10}\)The parameter \( \alpha \) allows me to study what happens when we shift some uncertainty from the forecastable factors \( z \) to the unforecastable \( \varepsilon \), while keeping the total amount of uncertainty \( \sigma^2_\gamma = \sigma^2_\zeta + \sigma^2_\varepsilon \) fixed.
asymmetry increases with risk aversion, and in particular, one can see that increasing risk aversion has three different effects. First, it increases the value of \( \kappa \) at which the tipping point occurs. Second, it increases the information asymmetry at any given \( \kappa \). And third, it increases the value to which information asymmetry converges as capacity is made limitless. The last result is not mentioned explicitly in Proposition 4, but it can easily be verified from the proof of Proposition 3.

Figure 2: Information Asymmetry as a Function of Risk Aversion

Figure 3 plots the same quantities, but considers changes in \( \delta \) instead of \( \gamma \). The resulting picture is very similar - \( \delta \) varies both the tipping point and the amount of information asymmetry for any given \( \kappa \). Increasing \( \delta \) increases the share of labor income in an agent’s total income, making labor income risk a larger concern. Since non-diversifiable labor risk is what drives the information asymmetry, it is not surprising to see that the skew in attention allocation increases as \( \delta \) (and hence the labor/non-financial share of income) increases.

Lastly, Figure 4 presents plots for three different values of \( \alpha \). One can see
that the attention tipping point $\bar{\kappa}$ increases as $\alpha$ increases - this is because the incentives for specialization increase as the learnable information can now reduce a larger fraction of the total uncertainty. However, unlike $\gamma$ and $\delta$, $\alpha$ also has another effect. Increasing $\alpha$ decreases the terminal value of information asymmetry because $\alpha$ also controls the correlation between the home asset and labor income. A higher $\alpha$ implies a higher correlation and makes the home asset a better hedge for labor income. In turn, this results in a lower exposure to undiversifiable labor income risk, which then makes home information relatively less valuable. This second effect dominates for high values of $\kappa$, and hence we observe the crossing of the lines in Figure 4. This result is also not explicitly mentioned in Proposition 4, but can be easily verified from the expressions for the constants $A$ and $B$ cited in Proposition 3.\footnote{In any case, Proposition 1 holds, and some information asymmetry always exists.}

The effect of the relative size of unlearnable to learnable uncertainty on the
Figure 4: Information Asymmetry and Unlearnable Uncertainty

information acquisition choice is quite intuitive in its own right. As with everything else, specialization in information acquisition is only appealing when additional information on the same topic could potentially lead to a relatively high payoff. If \( \alpha \) is close to zero, even perfect information about either \( z_h \) or \( z_f \) offers little reward. If \( \alpha \) is close to one, however, precise information about one of the forecastable factors is very valuable to the individual, as it comes very close to perfectly revealing the future.

5.4 Endogenous vs Exogenous Information Asymmetry

Van Nieuwerburgh and Veldkamp (2009, 2010) develop a framework in which endogenous information acquisition generates information asymmetry by amplifying any prior exogenous information advantages. The exogenously given prior information advantage is key, as the model will not generate any information asymmetry without it, but it can be arbitrarily small. This is a strong and important result but
unfortunately does not hold in a setup with unlearnable uncertainty ($\sigma_\varepsilon^2 > 0$). This section provides details on this argument and shows that a theory of endogenous information asymmetry is necessary for a robust information-based explanation of portfolio under-diversification.

The model in Van Nieuwerburgh and Veldkamp (2009) and in Corollary 2 of Van Nieuwerburgh and Veldkamp (2010) is a special case of the model presented in this paper, which is obtained by setting $\delta = 0$ and $\sigma_\varepsilon^2 = 0$. In other words, the framework of this paper adds the concept of unlearnable uncertainty and labor income to the previous literature’s work. I will refer to the restricted model as the “benchmark exogenous information asymmetry” model. I classify it as a model of exogenous information asymmetry, because even though agents make optimal information acquisition decisions, the results rely on the assumption that home agents have prior information advantages over the home assets.

In an environment with no labor income ($\delta = 0$) and no unlearnable uncertainty ($\sigma_\varepsilon^2 = 0$) the model exhibits increasing returns to information acquisition over the whole parameter space. Thus, the optimal strategy of the agents is to always exhaust their whole information processing capacity on one and only one of the forecastable factors. Which of the two assets the agent will choose to specialize in depends on the differences in his prior beliefs. If one endows the agent with symmetric prior information, as this paper does, then the solution is indeterminate - the agent will be indifferent between focusing only on the home factor or focusing only on the foreign factor. But if we endow the agents with even just an arbitrarily small information advantage in one of the factors, then the unique optimal strategy is to pay attention only to that factor. This result can be then used to conclude that small prior information advantages are not only sustained, but amplified through endogenous information acquisition.

This result does not hold in a more general setting where $\sigma_\varepsilon^2 > 0$ as this introduces decreasing returns to information acquisition. In particular, it is the case that the objective function will become concave when $\kappa > \bar{\kappa}$. At that instant, the agent’s optimal strategy changes so that he pays attention to both risk factors and he does so in such a way as to undo any and all prior information advantages. Given a large enough capacity, any prior information advantage will be eliminated.
and not amplified. Figure 5 shows how the size of information asymmetry behaves in the benchmark exogenous information asymmetry model, in a model that simply adds unlearnable uncertainty and finally the full model of this paper.

Figure 5: Models of Exogenous Information Asymmetry

The blue dashed line shows the standard Van Nieuwerburgh and Veldkamp (2010) framework, the red dash-dot line augments that model with unlearnable uncertainty and the solid green line is the model developed here. When the agent has no capacity for processing information (\(\kappa = 0\)) \(\Lambda\) is trivially zero in all models, and then it increases linearly with \(\kappa\). This is due to the increasing returns to information acquisition that would exist in all models for small values of \(\kappa\). However, if \(\sigma_\varepsilon^2 > 0\), once \(\kappa\) becomes bigger than \(\bar{\kappa}\) the optimal attention allocation strategy changes abruptly. Instead of only acquiring information about the home forecastable factor, the agents find it optimal to learn about both home and foreign factors, and in fact

---

\[12\] This result also holds for models that extend the Van Nieuwerburgh and Veldkamp (2010) framework along the lines of Mondria (2010). Such models allow the agents to observe an optimal linear combination of the two factors, rather than separate signals for each.
in the model with no labor income ($\delta = 0$) agents end up focusing most of their attention on the foreign factor. This is because the agents face a concave problem and the optimal solution is to have equally precise posteriors about both the home and foreign fundamentals. But to get there they first need to undo the exogenously assumed prior information asymmetry, and hence spend most of their information capacity on the foreign factor. The second graph makes this clear by plotting the ratio of posterior standard deviations $\frac{\sigma_f}{\sigma_h}$ as a function of $\kappa$ - notice the jump down to 1 for $\kappa > \bar{\kappa}$. The following proposition formalizes these results and more.

**Proposition 5.** Let $\delta = 0$ and the agents be endowed with a free signal about the home fundamentals, $s = z_h + u$, where $u \sim N(0, \sigma_u^2)$. Then,

- If $\sigma_e^2 = 0$, agents choose to devote their whole attention to the home factor, and hence we obtain $\Lambda = \kappa$.
- If $\sigma_e^2 > 0$, then there exists a $\bar{\kappa}(\sigma_e^2, \sigma_u^2)$ such that agents allocate their whole attention to the home factor when $\kappa < \bar{\kappa}$ but focus most of their attention on the foreign factor when $\kappa > \bar{\kappa}$. When $\kappa < \bar{\kappa}$, $\Lambda = \kappa$ and $\frac{\sigma_f}{\sigma_h} = \exp(2\kappa)$, and when $\kappa > \bar{\kappa}$, $\Lambda < 0$ and $\frac{\sigma_f}{\sigma_h} = 1$. Also, $\bar{\kappa}(\sigma_e^2, \sigma_u^2)$ is a decreasing function of $\sigma_e^2$ and an increasing function of $\sigma_u^2$.

**Proof.** The proof is in the Appendix.\[\square\]

Proposition 5 formalizes the intuition displayed in Figure 5. In the presence of unlearnable uncertainty, agents exploit the benefits of increasing returns to information acquisition and only allocate attention to the home factor when the information capacity constraint is tight. However, if the agents possess sufficient information processing capacity, they exhaust the increasing returns to information and instead prefer to be equally well informed about both factors. To do so, they need to unravel the prior information advantage they have over the home factor, and thus for high values of the capacity constraint the agents always pay more attention to the foreign factor rather than the home one.\[13\]

\[13\]In fact these results are even stronger if we generalize the information structure to allow the agents to observe a linear combination of the factors as in Mondria (2010). In that setting, agents acquire information more efficiently and in fact display a foreign bias in their information acquisition for all $\kappa$.\[\]
The last point in Proposition 5 states that the tipping point $\kappa$ is an increasing function of $\sigma_h^2$. The reason is that larger prior information advantage (lower $\sigma_h^2$) implies that the agents need to expend less of their capacity before they reduce the posterior uncertainty $\sigma_h^2$ to the point where they lose the increasing returns to information acquisition. In particular, this implies that if agents start with a sufficiently large prior information advantage the interval of increasing returns to information acquisition would not exist at all, and the agents will immediately undo it by paying most of their attention to the foreign fundamentals.

This result connects to one of the main criticisms of information asymmetry explanations of portfolio concentration – why agents do not learn about what they are most uncertain about. Non-diversifiable labor income can provide an answer, as it gives the agents an endogenous reason to value home information more than foreign information. The solid green line in Figure 5 represents the model with labor income. The graph shows it can generate home bias in information acquisition even for high values of $\kappa$, where the other model implies the opposite - a foreign bias. Most importantly, whereas the model without labor income implies that $\sigma_h^2 = \sigma_f^2$ for high values of $\kappa$, the model of this paper always produces $\sigma_h^2 < \sigma_f^2$. This latter result is crucial to explaining portfolio concentration, because in these frameworks the agents would only bias their portfolios towards one of the assets if they perceive it to be less risky, i.e. one needs $\sigma_h^2 < \sigma_f^2$.

Thus a model based solely on exogenous information asymmetry is not able to generate portfolio home bias when $\sigma_e^2 > 0$ and $\kappa > \bar{\kappa}$. More importantly, in Section 6 I present empirical evidence that the most likely description of the real world is indeed $\kappa > \bar{\kappa}$. This should not be surprising, because finding otherwise would suggest that investors enjoy increasing returns to information acquisition. But imagine a straightforward extension of the model, where instead of keeping $\kappa$ as a fixed parameter we allow agents to choose their optimal $\kappa$ subject to a simple, linear information cost function. In that case, the agents would never choose a $\kappa < \bar{\kappa}$ because it is optimal to exhaust all available increasing returns to information. The empirical findings in Section 6 confirm this intuition. Hence, a theory of endogenous information asymmetry is necessary for an information based explanation of portfolio concentration.
5.5 Quantitative Effect on Portfolio Allocations

This section analyzes the quantitative dimensions of the model and asks the question of how much portfolio concentration it can generate under a few reasonable calibrations.

I compare the model against three benchmark models of standard Bayesian agents. I refer to the first one as “Equal Information” because the agents in it receive equally informative signals with a total information content of $\kappa$. I also consider two models of exogenous information asymmetry. In the first model the agents are assumed to receive a home signal with 25% higher precision than their foreign signal, and in the second the home signal is twice as precise as the foreign signal.

For this numerical exercise, I follow Van Nieuwerburgh and Veldkamp (2009) and set the mean payoff of the assets to 1 and the standard deviation to 15%, and following Mondria (2010) I set the gross return of the riskless bond to 1.02 and the coefficient of absolute risk aversion $\gamma$ to 2. I choose $\delta$ so that on average the agents’ labor share is 0.68, a standard value in the RBC literature.14 Since the previous literature has not considered unlearnable uncertainty, there is little guidance for the value of $\alpha$. In order to better illustrate the results of the model I will consider two different values: $\alpha = \frac{2}{3}$ and $\alpha = \frac{3}{4}$. I also calibrate two different values of $p_h = p_f$, one which implies a relatively small risk premium of 3% and another value which implies a risk premium of 6%.

Given all other parameters, I calibrate $\kappa$ so that the model implies a set amount of attention is allocated to the foreign factor. In particular, I consider values for $\kappa$ such that the agent allocates 5%, 15%, 25%, 35% and 45% of available capacity to the foreign factor. This range of values was chosen to cover the whole

---

14 The average investor is wealthier than the average person, and hence at first glance this $\delta$ may appear poorly chosen, but this is not the case. The 2007 Survey of Consumer Finances (SCF) shows that people in the top decile for wealth get only about 50% of their income from wages, but do get about 20% from private business and self-employment, which is another source of non-financial income. Counting both wages and self-employment income as non-financial income, the SCF suggests that non-financial income accounts for about 70 – 80% of the total income for all wealth groups. Hence a non-financial income share of 0.68 is consistent with the empirical evidence.
The results in Table 1 indicate that the model can deliver sizable amounts of portfolio home bias for a large number of parameters. The results are strongest when the risk premium is high (6%) and the amount of unlearnable uncertainty is low ($\alpha = \frac{3}{4}$). In that case, the model delivers home bias in portfolios even when the attention of the agent is split fairly evenly among home and foreign matters, and the portfolio share of home assets rises above 60% if agents allocate 25% or less of their attention to the foreign fundamental. The results are somewhat weakened when the risk premium decreases to 3% - in this case, home bias arises if agents...
pay less than 25% of their attention to the foreign fundamental and home assets reach above 60% of the total portfolio once the foreign share of attention dips below 15%. On the other hand, the generated home bias is also smaller if we increase the amount of unlearnable uncertainty as this decreases the incentive for specialization in information acquisition. The model can still generate considerable amounts of home bias but in all cases the values are lower than in the case of small unlearnable uncertainty. In particular, the model is most challenged under a parameterization where both the risk premium is low and the unlearnable uncertainty is large. In this case, one needs to calibrate $\kappa$ so that the agent allocates less than 5% of his attention to the foreign assets in order to get a considerable amount of portfolio home bias.

The main takeaway of the table, however, is that the framework performs very well relative to the three benchmark models. It generates significantly higher amounts of home bias under most parameterizations. Perhaps even more importantly, it generates at least some home bias under most parameterizations, while the other models virtually always imply a foreign bias. The foreign bias arises because of the strong hedging incentive to sell the home asset due to its positive correlation with labor income. This paper overcomes this strong force by allowing the agents to optimally choose their information acquisition strategy. In this setting, the labor income risk does not only compel agents to short the home asset for hedging purposes, but also makes home information more valuable, and in turn generates information asymmetry which provides a positive demand for the home asset. The two forces act opposite to each other, but the table demonstrates that the information channel dominates the hedging incentive. On the other hand, the table makes clear that the models of exogenous information asymmetry are not able to overturn the hedging incentives, even though they assume significant information advantages over the home asset.

Lastly, Figure 6 puts these numbers in context by comparing them to the data. It uses the values in Table 1 to compute and graph the implied Equity Home Bias index (EHB) for the model of this paper and the strong exogenous information advantage model. The EHB index is a commonly used measure of portfolio concentration in the literature on home equity bias. It is defined as,
\[ EHB = 1 - \frac{\text{Foreign Equity as a Share of the National Portfolio}}{\text{Foreign Equity as a Share of the World Market Portfolio}} \]

and measures how much a country’s equity portfolio deviates from the global market portfolio.\(^{17}\) The index is positive when the analyzed portfolio differs from the market portfolio by holding a larger proportion of home assets (i.e. exhibits home bias), negative when it exhibits foreign bias, and is 0 when it is exactly equal to the market portfolio. In addition to the EHB index implied by the values in Table 1, the figure also plots two horizontal lines which are the average EHB found in international data and the average EHB found in domestic data.\(^{18}\)

Figure 6: Exogenous vs Endogenous Information Asymmetry

\(^{17}\)The market portfolio is the equilibrium portfolio in standard CAPM theory (Sharpe (1964)) and is the most commonly used benchmark of the “fully diversified” portfolio.

\(^{18}\)The sources of the international data are described in Section 6. The domestic EHB is calculated from data presented in Ivković and Weisbenner (2005) by classifying holdings of “local” firms as “home equity assets”, and “non-local” firms as “foreign equity assets”. In the case of domestic data, the index then measures how much is a portfolio biased towards local stocks, compared to the domestic market portfolio.
The graph shows that the theoretical framework in this paper can generate home bias in the neighborhood of the estimates found in domestic data, but falls short of the high values estimated from international data. This is most likely due to the fact that international financial markets have a lot more frictions than modeled here. The frictionless environment of the paper is likely a much better description of the domestic market and thus it is encouraging to see that the model can match the domestic data easily. In any case, the endogenous information asymmetry model represents a vast improvement over the benchmark model of exogenous information asymmetry.\textsuperscript{19}

6 Empirical Evidence

Section 5.4 shows that under the parametric restriction of $\kappa < \bar{k}$, the exogenous information advantage models in the spirit of Van Nieuwerburgh and Veldkamp (2009, 2010) deliver the same results and implications as the endogenous information asymmetry model developed in this paper. In light of this, it is fair to ask if a model of endogenous information asymmetry is empirically relevant and if such theory is needed in order to fully understand the phenomenon of portfolio concentration. Here I address this question by exploiting the fact that whenever $\kappa \geq \bar{k}$, the two frameworks have markedly different implications for relationship between portfolio concentration, the level of information capacity, and the relative size of non-financial income.\textsuperscript{20}

\textsuperscript{19}A criticism of the literature on information asymmetry and the home bias is that it often abstracts from the existence of investment intermediaries (e.g. mutual funds). At first glance, it seems that mutual funds could be well informed about both home and foreign events and offer the agents a well-diversified investment opportunity. However, introducing mutual funds is unlikely to change the main results of the paper because of issues such as principal-agent problems and uncertain fund quality. In Appendix B I discuss some reasons why an extension along these lines would not have major effects on the paper’s results, but a detailed analysis is left to future work.

\textsuperscript{20}Another clear difference between the models arises in an N-asset framework. In an exogenous information advantage model, the agents only acquire information about the home asset and thus have symmetric information over all foreign assets and hold all foreign assets in the same proportion. On the other hand, in an endogenous information asymmetry model, the agents stagger foreign information acquisition and thus for any $\kappa$, they possess information of different quality for the different foreign assets, which leaves them holding the foreign assets in different proportions. Empirical evidence shows that countries do not treat all foreign assets as equal, and the composition of the foreign equity portfolio varies greatly across countries. This empirical regularity could also be used to differentiate the two models but the formal treatment is left to
The empirical restrictions examined in this section are derived from Propositions 3, 4 and 5. Proposition 5 shows that in exogenous information advantage models, agents exhaust their entire information acquisition ability on the home asset, and hence the implied information asymmetry is increasing in the capacity constraint $\kappa$. This paper’s framework, however, has a different prediction. According to Proposition 3, the magnitude of the information asymmetry decreases with $\kappa$, whenever $\kappa \geq \bar{\kappa}$. This suggests that one way to differentiate the two frameworks is to look at the empirical relationship between portfolio bias and information processing constraints – a negative relationship will be at odds with the exogenous information advantage models but consistent with the framework of this paper.

Moreover, in the exogenous information advantage models, the relative size of labor income does not affect the information acquisition decisions of the agents. In such models labor income has only one effect on portfolio choice - it provides a hedging motive that compels the agents to tilt their portfolios towards foreign assets. In the model of this paper, however, labor income does affect the information acquisition decisions of the agents, and as a result, agents with a high labor income share use more of their information capacity on home fundamentals and less of it on foreign fundamentals (Proposition 4). In this setting, labor income has a second effect on portfolios, which operates through its effect on information acquisition, and this effect is opposite in sign to the hedging motive. Hence, a positive empirical relationship can be explained by the model presented in this paper, but not by exogenous information asymmetry models.\(^21\) This section examines these relationships by regressing the Equity Home Bias index (EHB) on a measure of information capacity, labor income, and financial wealth, together with additional controls. The findings reject the implication of the exogenous information advantage model and are consistent with the implications of the model of this paper.

To implement this regression I have obtained aggregate national data on 35 OECD countries for the time period between 2001 and 2008. I have data on

---

\(^21\) Coeurdacier and Gourinchas (2011) also develop a model in which labor income generates home bias. Unfortunately, the aggregate data cannot distinguish between the mechanisms of this paper and Coeurdacier and Gourinchas (2011). However, micro data on domestic portfolios can discriminate between the two because the mechanism in Coeurdacier and Gourinchas (2011) relies on the existence of real exchange rate risk, which is absent in a domestic setting. At the end of the section, I will revisit this point using micro evidence from Massa and Simonov (2006).
portfolio positions, proxies for $\kappa$, labor income and financial wealth per capita, the Chinn-Ito index of financial openness and real PPP GDP per capita. The data comes from the IMF, the OECD and the World Bank and full details are given in the online Appendix.

As a proxy for information capacity, $\kappa$, I use the number of Internet users per 100 people, which I obtain from the World Bank. The availability of the Internet is a natural proxy for information processing capacity for two reasons. First, Internet technology greatly reduces the time associated with collecting information, and second, an Internet connection also presupposes the availability of a computer, which further enhances a person’s ability to sift through and analyze information. It is straightforward that for any fixed period of time, an Internet user can process more bits of information than a person without an Internet connection. Lastly, for the time period at hand - 2001 through 2008 - the Internet was readily available to all 35 countries in the sample, but at the same time the countries exhibit significant variation in the extent to which the Internet had penetrated their populations.\textsuperscript{22}

Next, I average all variables across time to arrive at a cross-section with 35 observations, and I run the following regression, \textsuperscript{23}

$$EHB_i = const + \beta_k \ln(I_{neti}) + \beta_l \ln(LabIncome_i) + \beta_f \ln(FinWealth_i) + \beta X_i + \varepsilon_i$$

where the vector $X_i$ is a vector of control variables. The signs of $\beta_k$, $\beta_l$ and $\beta_f$ are of primary interest. The regression includes Financial Wealth, rather than

\textsuperscript{22}The Appendix also presents results for two alternative proxies for $\kappa$ - cell phones per capita and computers per capita. The coefficient on cell phone users is negative, and of the same magnitude as the one on Internet users, but it is not as efficiently estimated and as a result is significant only at the 17% level. On the other hand, the availability of computers per capita data is much worse and this drops the sample observations to just 24 and the regression appears to run into multicollinearity problems with a highly significant $R^2$ around 0.6 but only a single individually significant coefficient (with a t-stat of about 2). As a result it is perhaps not surprising that the estimated coefficients on PCs per capita are not insignificant. In any case, none of the different specifications finds a significant positive coefficient.

\textsuperscript{23}I focus my attention on time averages because both models are static and have nothing to say about dynamic portfolio choice decisions. Nevertheless, panel regressions also find a negative and statistically significant relationship between Internet Users per 100 ppl and home bias, and a statistically insignificant relationship between home bias and labor income and financial wealth. The results are not reported here because the testable restrictions are derived from static models and it is unclear what are the relevant null and alternative hypotheses in a dynamic setting, but the results are available upon request.
Financial Income, because of the lack of data on the latter. Financial Wealth does not include information on the rate of return earned by different countries on their portfolios and if there is significant cross-country variation in those rates it could bias the estimated coefficient. However, in the Appendix I also report estimates with a couple of different formulations and/or variables and find the results to be virtually unchanged. The vector of controls includes the Chinn-Ito Index, which measures the degree of existing capital controls in a given country, and the initial period (year 2001) level of real PPP GDP per capita. The first variable controls for the possibility that home bias is a result of international capital flows frictions. The second controls for initial conditions, and especially ensures that the coefficient on Internet Users per 100 people does not simply pick up the fact that rich countries tend to be more diversified.

The first empirical restriction I test can be summarized by $H_0 : \beta_k > 0$ vs $H_a : \beta_k \leq 0$, and the second is given by $H_0 : \beta_1 < 0, \beta_f > 0$ vs $H_a : \beta_1 \geq 0, \beta_f \leq 0$. A rejection of one of the null hypothesis would mean that the data rejects an implication of the exogenous information advantage model, but is consistent with the endogenous information asymmetry model. Technically both hypotheses are one-sided, however, to be conservative all results are reported at the standard 2-sided confidence levels.

Table 2 presents OLS estimates with the corresponding heteroskedasticity robust standard errors in parentheses. The first column shows that the unconditional correlation between Internet users (the measure for information capacity) and home

---

24 The online Appendix shows that the results are unchanged if we impose the restriction that $\beta_1 = \beta_f$, i.e. regressing on the ratio (however, the restriction is rejected by the data (p-value of 0.02)). The Appendix also shows that the results are unchanged if I use the standard quantity of labor income share of GDP instead.

25 Stockman and Tesar (1995) and others have shown that non-tradable consumption can also be responsible for the equity home bias puzzle. The results are very similar when controlling for the non-tradables and are reported in the online Appendix. I do not include non-tradables in the main body of the paper for two reasons. First, such data is scarcer and it drops the sample to only 24 countries and, second, the previous literature (Pesenti and Van Wincoop (2002), Lewis (1999)) has shown that non-tradables are unlikely to be important determinants of the home bias.

26 The online Appendix includes results from a regression where I include the square of $\kappa$ to capture potential non-linearities (which is, strictly speaking, the model’s prediction). This more flexible setup does find the non-monotonic relationship predicted by Proposition 3 - EHB is increasing in $\kappa$ for countries with low $\kappa$ and decreasing otherwise (only 5 countries are found to be in the increasing region). I relegate these results to the Appendix in order to keep the main regression specification as parsimonious as possible.
bias is significantly negative. The second and the third columns add the measures of per capita Labor Income and Financial Wealth, the fourth and fifth add the Chinn-Ito Index of financial openness and the last two columns present results including the last control variable, per capita PPP RGDP. The table reports results for two different proxies for financial wealth - Total Equity Holdings and Total Financial Assets.\textsuperscript{27}

Under all specifications, the number of Internet users per 100 people has a statistically significant negative relationship with the degree of home bias. Moreover, the point estimate is roughly the same through columns (2)-(7) and also throughout the alternative specifications reported in the Appendix, which suggests the estimate is robust. Thus, the data rejects the first empirical restriction.\textsuperscript{28} The coefficient

\begin{table}[h]
\centering
\caption{Regression Results}
\begin{tabular}{lccccccc}
\hline
 & (1) & (2) & (3) & (4) & (5) & (6) & (7) \\
\hline
Log Internet Users & -0.236*** & -0.106* & -0.101* & -0.095* & -0.099* & -0.113** & -0.114** \\
 & (0.029) & (0.055) & (0.059) & (0.057) & (0.059) & (0.049) & (0.049) \\
Log Labor Income & -0.132** & -0.182 & -0.033 & -0.045 & 0.259** & 0.257** \\
 & (0.065) & (0.109) & (0.062) & (0.119) & (0.101) & (0.119) \\
Log Total Fin Assets & -0.019 & -0.035* & -0.019 \\
 & (0.024) & (0.019) & (0.015) \\
Log Equity Assets & 0.006 & -0.024 & 0.001 \\
 & (0.034) & (0.036) & (0.032) \\
Chinn-Ito Index & -0.065** & -0.059** & -0.060*** & -0.055*** \\
 & (0.024) & (0.025) & (0.018) & (0.017) \\
Log RGDP\textsubscript{2001} p.c. & -0.327*** & -0.366*** \\
 & (0.095) & (0.093) \\
\hline
N & 35 & 35 & 35 & 35 & 35 & 35 & 35 \\
R\textsuperscript{2} & 0.475 & 0.568 & 0.561 & 0.638 & 0.618 & 0.684 & 0.678 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{27}Total Financial Assets is more appealing theoretically, but I do not have direct observations on this measure and impute it from other available data. Thus, I also present results with Total Equity Assets which is directly observable, but is not a complete measure of Financial Wealth. Details on how Total Financial Assets and other variables are constructed are given in the Appendix.

\textsuperscript{28}Mondria and Wu (2010) use a slightly different proxy for $\kappa$, which is a measure of IT technology per $1000$ of economic activity rather than IT technology per capita, and find the opposite result that the Home Bias is positively related with their measure of $\kappa$. In the Appendix I reestimate my regressions using the Mondria and Wu (2010) measure of $\kappa$ and again find a statistically significant negative relationship. There I also discuss three possible reasons for the difference in the results: the different data sets, the extra controls I include, and differences in methodology.
on Labor Income in the full model (the last two columns) is found to be positive and statistically significant, while the coefficients on the two measures of financial wealth are insignificant. Nevertheless, the coefficient on Total Financial Assets in column (6) is marginally significant in a one-sided test (p-value 0.1), which is suggestive of the expected negative relationship. Hence, by virtue of finding a significant positive coefficient on Labor Income and an insignificant coefficient on Financial Wealth, I conclude that the data also rejects the second empirical restriction.

Lastly, note that the coefficient on Labor Income is negative and significant in columns 2 and 3, and insignificant in columns 3 and 4. This appears to be due to Labor Income per capita being highly correlated with the Chinn-Ito index of financial openness and GDP per capita. Once we control for the degree of financial openness and the relative richness/poorness of the country, the estimate is found to be positive. Even without controlling for these additional effects, the coefficient on Financial Wealth is still estimated to be negative, and is marginally significant once we control for the degree of financial openness. The results across all columns are consistent with the rejection of the second empirical restriction.

Interestingly, this is a phenomenon present in micro-level data as well. Using data on each individual investor’s labor income, financial income and portfolio composition, Massa and Simonov (2006) find that investors actively tilt their portfolios towards assets that are positively correlated with their labor income process. They find that while the average correlation between the stock market as a whole and an individual’s labor income is roughly zero, the average correlation between an individual’s labor income and his portfolio proceeds is positive and statistically significant. Moreover, they also find evidence that investors act deliberately in skewing their portfolios towards assets that are positively correlated with their labor incomes. They show on average the stocks that investors buy increases the correlation of portfolio returns and non-financial income, while the stocks they sell decrease this correlation. Most importantly, they find that the size of portfolio bias is negatively related to the investor’s total wealth. They find that on average, the bias amounts to 41% of the total risky asset holdings of low wealth investors and for 10% of the risky assets of wealthy investors. This suggests that on the micro level, investors that are more dependent on non-financial income have more
concetrated portfolios, and this is again evidence against the second empirical restriction again.\(^29\)

Thus, this is a robust pattern found both in micro and macro data. Coeurdacier and Gourinchas (2011) develop an open economy macroeconomic model where the presence of real exchange rate uncertainty and non-tradable labor income can also deliver the result that home bias is positively related to the size of labor income. However, that mechanism will have difficulty matching the micro level data. The model relies on three things: 1) real exchange rate risk, 2) the availability of a local-currency denominated bond with returns that co-move strongly with the local price level and 3) equity returns which are negatively related with labor income, conditional on the bond returns. Under this setup, the agents choose to use bonds to hedge the exchange rate risk and then take on a home biased position in equities to hedge their remaining labor income risk. Coeurdacier and Gourinchas (2011) and others show that the necessary conditions are generally satisfied by the joint distribution of international price levels, bonds and equities. It is harder, however, to imagine that such locality-tied nominal bond instruments exist inside a single country. At the same time, the mechanism of this paper operates both inside a single country and on the international level. It provides an explanation of the general phenomenon of portfolio concentration, and is not particular to domestic or international market conditions.\(^30\)

Lastly, note that this paper considers only technological measures of information processing capacity. While in the 21\textsuperscript{st} century information technology is certainly a very important, if not the chief, determinant of information processing ability, human capital is also likely to play an important role. Examining the empirical relationship between portfolio concentration and investors’ human capital is beyond the scope of this paper, but there already exists an empirical literature on this issue. The interested reader is directed to Goetzmann and Kumar (2008) who show that portfolio concentration is decreasing in investors’ educational level, investing experience and sophistication, and also to, Kimball and Shumway (2010) who document that the international home bias is decreasing in a number of different

\(^{29}\)This is also in accordance with a large part of the empirical literature on under-diversification which documents that wealthier investors hold better diversified portfolios.

\(^{30}\)It is important to also note that the two mechanisms are not rival and in fact will amplify each other’s results if put together.
measures of investor sophistication.

In conclusion, this section set out to use data to differentiate between the model derived in this paper and the exogenous information advantage model used in previous work. The empirical evidence rejects two key implications of the exogenous information asymmetry model and is consistent instead with an endogenous information asymmetry framework. This suggests that there are empirically relevant differences in the two frameworks, and hence endogenous information asymmetry is likely important for a complete understanding of how information imperfections affect the phenomenon of portfolio concentration.

7 Conclusion

This paper addresses one of the major puzzles to financial and international economics - the systematic difference between theoretically optimal portfolios and portfolios observed in the data. It develops a framework in which agents who start with identical information over all assets optimally choose to use their limited ability to acquire information to learn mostly about assets which are closely related to their non-financial income. This results in ex-post information asymmetry which compels the agents to hold portfolios that are “biased” from the viewpoint of traditional portfolio choice theory. In particular, the agents choose portfolio allocations that are biased towards assets exposed to the same risk factors as their labor income, which offers an explanation of the puzzling empirical phenomenon that investors tend to hold portfolios concentrated in assets that are positively correlated with their non-financial income. An important implication of the model is that such biases arises as the optimal choice of rational agents, and are not due to incorrect beliefs, behavioral biases or unsophistication. The portfolios will appear inefficient to a third party analyst who does not observe the information sets of the individual investors, yet they are optimal.

It was shown that exogenous information advantage models have counterfactual implications which can be avoided when we move to a framework of endogenous information asymmetry. Thus, a key message of the paper is that a theory of endogenous information asymmetry appears to play an important role in information-based explanations of portfolio concentration. Here I showed that
non-financial income considerations are one way to generate information asymmetry endogenously but there are surely many more. The important thing is to think about channels through which information about asset returns also has an effect on the other decisions of the agents (e.g. housing choice, labor choice, etc.). Studying such additional channels in detail appears to be a fruitful avenue for future work.

References


A Appendix A: Proofs

A.1 Proposition 1:

Proof. By evaluating the log-normal expectation conditional on time 1 information and taking the log of the resulting expression I arrive at a non-central chi-square random variable. Then, using the formula for the expectation of chi-square variables, the agent’s maximization problem becomes:

$$\max_{\sigma_h^2, \sigma_f^2} U(\sigma_h^2, \sigma_f^2) = \gamma \delta \mu + \frac{1}{2}[-2 + \frac{\sigma_h^2 + \sigma_e^2}{\sigma_h^2 + \sigma_e^2}(1 + \frac{(\mu - p_h R)^2}{\sigma_e^2 + \sigma_e^2}(1 + \frac{(\mu - p_f R)^2}{\sigma_e^2 + \sigma_e^2})]$$

$$- \gamma \delta (\mu - p_h R) \frac{\sigma_h^2}{\sigma_h^2 + \sigma_e^2} - \gamma^2 \delta^2 \frac{\sigma_e^2}{2} \sigma_e^2 - \gamma^2 \delta^2 \frac{\sigma_e^2 \sigma_f^2}{2} \sigma_f^2$$

s.t.

$$\frac{1}{2} (\ln(\sigma^2_z) - \ln(\sigma^2_h) - \ln(\sigma^2_f)) \leq \kappa$$

$$-\infty \leq \ln(\sigma^2_z) \leq \ln(\sigma^2_h), \quad -\infty \leq \ln(\sigma^2_f) \leq \ln(\sigma^2_z)$$

The objective function is continuous and the domain of maximization is compact hence a maximum always exists. Since all constraints are linear in the log of the posterior variances it is more convenient to find the maximum in terms of the variables $\ln(\sigma^2_h)$ and $\ln(\sigma^2_f)$, rather than the posterior variances $\sigma^2_h$ and $\sigma^2_f$ themselves. Thus, to find the maximum reparameterize the problem in terms of $\ln(\sigma^2_h)$ and $\ln(\sigma^2_f)$ and setup the Lagrangian:

$$L = U(\exp(\ln(\sigma^2_h)), \exp(\ln(\sigma^2_f))) + \lambda (2\kappa - \ln(\sigma^2_z) + \ln(\sigma^2_h) + \ln(\sigma^2_f)) + \psi_h (\ln(\sigma^2_z) - \ln(\sigma^2_h)) + \psi_f (\ln(\sigma^2_z) - \ln(\sigma^2_f))$$

The Karush-Kuhn-Tucker necessary optimality conditions are:

$$\frac{\partial U}{\partial \ln(\sigma^2_h)} (\exp(\ln(\sigma^2_h)), \exp(\ln(\sigma^2_f))) = -\lambda + \psi_h$$

$$\frac{\partial U}{\partial \ln(\sigma^2_f)} (\exp(\ln(\sigma^2_h)), \exp(\ln(\sigma^2_f))) = -\lambda + \psi_f$$

$$(2\kappa - \ln(\sigma^2_z) + \ln(\sigma^2_h) + \ln(\sigma^2_f)) \lambda = 0$$

$$\psi_h (\ln(\sigma^2_z) - \ln(\sigma^2_h)) = 0$$

$$\psi_f (\ln(\sigma^2_z) - \ln(\sigma^2_f)) = 0$$

$$\lambda \geq 0, \quad \psi_h \geq 0, \quad \psi_f \geq 0$$
Before we proceed, I compute the first derivatives of the utility function for future reference:

\[
\frac{\partial U}{\partial \ln(\sigma_h^2)} = A \frac{\sigma_h^2}{(\sigma_h^2 + \sigma_{zh}^2)^2}
\]

\[
\frac{\partial U}{\partial \ln(\sigma_f^2)} = B \frac{\sigma_f^2}{(\sigma_f^2 + \sigma_{zf}^2)^2}
\]

where

\[
A = -\left[ \frac{1}{2} \left( (\sigma_z^2 + \sigma_{zh}^2) + \frac{(\mu - p_h R)^2}{\sigma_z^2 + \sigma_{zh}^2} \right) + \gamma \delta (\mu - p_h R) \sigma_{zh}^2 + \frac{\gamma^2 \delta^2}{2} \sigma_{zh}^4 \right]
\]

\[
B = -\left[ \frac{1}{2} \left( (\sigma_z^2 + \sigma_{zf}^2) + \frac{(\mu - p_f R)^2}{\sigma_z^2 + \sigma_{zf}^2} \right) \right]
\]

Next, let’s analyze the KKT necessary conditions. There are eight cases to consider.

Case I:

\[ \lambda = \psi_h = \psi_f = 0 \]

This is the case when all constraints are lax and the KKT conditions imply that at the maximum we have

\[
\frac{\partial U}{\partial \ln(\sigma_h^2)} = \frac{\partial U}{\partial \ln(\sigma_f^2)} = 0
\]

This is only possible if \( \sigma_h^2 = \sigma_f^2 = 0 \) which will violate the entropy constraint because \( \ln(\sigma_f^2) - 2 \ln(0) = \infty > 2\kappa \). Hence, this case is infeasible.

Case II:

\[ \lambda = 0, \, \psi_h > 0, \, \psi_f > 0 \]

This is the case where the entropy constraint is lax, but the “no forgetting” constraints \( \ln(\sigma_h^2) \leq \ln(\sigma_z^2) \) and \( \ln(\sigma_f^2) \leq \ln(\sigma_z^2) \) are both binding. Hence, we must have \( \sigma_h^2 = \sigma_f^2 = \sigma_z^2 \) and the information capacity is not used at all. However, the derivatives of the utility function \( \frac{\partial U}{\partial \ln(\sigma_h^2)} \) and \( \frac{\partial U}{\partial \ln(\sigma_f^2)} \) are both strictly negative and thus the agent can acquire strictly greater utility by setting \( \sigma_h^2 < \sigma_z^2 \) or \( \sigma_f^2 < \sigma_z^2 \). Therefore, this case does not characterize the maximum of the problem either.

Case III:

\[ \lambda = 0, \, \psi_h = 0, \, \psi_f > 0 \]

In this case the “no forgetting” constraint on the foreign fundamental is binding, and the other two are lax. Hence \( \sigma_f^2 = \sigma_z^2 \) and because the partial derivative is always strictly negative, it is optimal to exhaust the whole information constraint and set \( \sigma_h^2 = \frac{\sigma_z^2}{\exp(2\kappa)} \).

For now note that this can indeed characterize the maximum and I will derive the specific conditions under which it does turn out to be the global optimum further below.

Case IV:

\[ \lambda = 0, \, \psi_h > 0, \, \psi_f = 0 \]
This is the mirror case, where the “no forgetting” constraint on the home fundamental is binding and hence \( \sigma_h^2 = \sigma_f^2 \) and \( \sigma_h^2 = \frac{\sigma_f^2}{\exp(2\kappa)} \). This allocation always achieves utility which is strictly lower than the utility achieved in Case III, hence it cannot be optimal. The key to this result is that \( A < B \) and hence

\[
\frac{\partial U}{\partial \ln(\sigma_h^2)}(\exp(\ln(\sigma^2)), x) < \frac{\partial U}{\partial \ln(\sigma_f^2)}(y, \exp(\ln(\sigma^2))), \forall x, y
\]

(1)

Thus, we have

\[
U\left(\frac{\sigma_h^2}{\exp(2\kappa)}, \sigma_f^2\right) - U\left(\frac{\sigma_f^2}{\exp(2\kappa)}, \frac{\sigma_h^2}{\exp(2\kappa)}\right) = \int_{\ln(\sigma_f^2)-2\kappa}^{\ln(\sigma_h^2)} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(y), \sigma_f^2)dy + \int_{\ln(\sigma_h^2)}^{\ln(\sigma_f^2)} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_f^2, \exp(y))dy
\]

\[
= \int_{\ln(\sigma_f^2)-2\kappa}^{\ln(\sigma_h^2)} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_f^2, \exp(y)) - \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(y), \sigma_f^2)dy
\]

\[
> 0
\]

The first equality follows from the Fundamental Theorem of Calculus and the third line inequality follows from (1). Thus, allocating all attention to \( z_h \) is always strictly better than allocating all attention to \( z_f \) and hence Case III dominates Case IV.

Case V: \( \lambda > 0, \psi_h = \psi_f = 0 \)

In this situation the information capacity constraint binds and the “no forgetting” constraints are lax. It follows that information acquisition decision satisfies

\[
\frac{\partial U}{\partial \ln(\sigma_h^2)} = \frac{\partial U}{\partial \ln(\sigma_f^2)} = -\lambda < 0
\]

\[
\ln(\sigma_h^4) - \ln(\sigma_f^4) - \ln(\sigma_f^2) = \kappa
\]

This can also characterize the maximum and the specific conditions will be characterized below. I will refer to this case as the “interior solution”.

Case VI: \( \lambda > 0, \psi_h > 0, \psi_f > 0 \)

This is a situation where all constraints are binding and this is impossible. Whenever \( \sigma_h^2 = \sigma_f^2 = \sigma_f^2 \) it follows that the information acquisition constraint does not bind, and thus one of the complimentary slackness conditions is not satisfied. This cannot be a maximum.

Case VII: \( \lambda > 0, \psi_h = 0, \psi_f > 0 \)

This is the case where the information capacity constraint and the “no forgetting” constraint on \( z_f \) are binding and the other “no forgetting” constraint is lax. This results in \( \sigma_f^2 = \sigma_f^2 \) and \( \sigma_h^2 = \frac{\sigma_f^2}{\exp(2\kappa)} \) which are the exact same allocations as in Case III - Case III and Case VII are identical.

Case VIII: \( \lambda > 0, \psi_h > 0, \psi_f = 0 \)

This case is identical to Case IV and for the same reason, it achieves strictly lower utility than Cases III and VII, hence it cannot be optimal.

Having analyzed all possible cases, it turns out that there are only two types of information acquisition allocations which could be optimal - \( \{\sigma_h^2, \sigma_f^2\} \) satisfy either:
\[
\frac{\partial U}{\partial \ln(\sigma_h^2)} = \frac{\partial U}{\partial \ln(\sigma_f^2)} = -\lambda < 0
\]

\[
\ln(\sigma_z^4) - \ln(\sigma_h^2) - \ln(\sigma_f^2) = 2\kappa
\]

OR

\[
\sigma_f^2 = \sigma_z^2, \quad \sigma_h^2 = \frac{\sigma_z^2}{\exp(2\kappa)}
\]

The first is an interior solution where the agent acquires information about both \(z_h\) and \(z_f\) and the second is a corner solution where the agent acquires information only about \(z_h\). Next, I analyze under what conditions we obtain one or the other.

First, I show that an allocation where \(\sigma_h^2 > \sigma_z^2\) and \(\sigma_f^2 < \sigma_z^2\) can never be optimal. In other words, unless the agent has enough capacity to drive \(\sigma_h^2\) at least as low as \(\sigma_z^2\), the corner solution is optimal and he will never allocate any attention to the foreign fundamental. The proof of this fact proceeds by contradiction and to this end let us assume that \(\sigma_h^2 > \sigma_z^2\) and \(\sigma_f^2 < \sigma_z^2\) obtains a global maximum of the utility function. This is an interior point (the corner solution where \(\sigma_h^2 = \sigma_z^2\) was already shown to never be optimal, so we only need to worry about possible interior solutions where \(\sigma_h^2 > \sigma_z^2\) and \(\sigma_f^2 < \sigma_z^2\)), hence the two partial derivatives must be equal to each other, i.e.:

\[
\frac{\partial U}{\partial \ln(\sigma_h^2)} = \frac{\partial U}{\partial \ln(\sigma_f^2)}
\]

Now, if we have \(\sigma_z^2 < \sigma_h^2\) and \(\sigma_z^2 < \sigma_f^2\) it follows from Proposition 2 (proved below) that both derivatives are increasing in their respective posterior variances \(\sigma_h^2\) and \(\sigma_f^2\) and thus there are increasing returns to information acquisition. Hence if we move \(\bar{\varepsilon}\) amount of attention from \(z_h\) to \(z_f\) so that we reduce \(\sigma_h^2\) down to \(\frac{\sigma_h^2}{\exp(\bar{\varepsilon})} \geq \sigma_z^2\) and simultaneously increase \(\sigma_f^2\) to \(\sigma_f^2 \exp(\bar{\varepsilon}) \leq \sigma_z^2\) (pick \(\bar{\varepsilon}\) small enough), the derivative on the left-hand side monotonically decreases, while the one on the right increases. Thus for all \(\varepsilon \in [0, \bar{\varepsilon}]\):

\[
\frac{\partial U}{\partial \ln(\sigma_h^2)}(\exp(\ln(\sigma^2) - \varepsilon), x) < \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(y, \ln(\sigma^2) + \varepsilon)), \forall x, y
\]

Therefore,

\[
U\left(\frac{\sigma_h^2}{\exp(\varepsilon)}, \sigma_f^2 \exp(\varepsilon)\right) - U(\sigma_h^2, \sigma_f^2) = \int_{\sigma_f^2}^{\sigma_h^2 \exp(\varepsilon)} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, x) d\ln(x) - \int_{\sigma_f^2}^{\sigma_h^2} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, \sigma_f^2 \exp(\varepsilon)) d\ln(x)
\]

\[
= \int_{\ln(\sigma_f^2)}^{\ln(\sigma_f^2) + \varepsilon} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, \exp(\varepsilon), \sigma_f^2 \exp(\varepsilon)) d\ln(x) - \int_{\ln(\sigma_f^2) - \varepsilon}^{\ln(\sigma_f^2)} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(y, \sigma_f^2 \exp(\varepsilon)) d\ln(x)
\]

\[
= \int_{\varepsilon}^{\bar{\varepsilon}} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, \exp(\ln(\sigma_f^2) + \varepsilon)) d\varepsilon - \int_{0}^{\varepsilon} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(\ln(\sigma_f^2) - \varepsilon), \sigma_f^2 \exp(\varepsilon)) d\varepsilon
\]

\[
> 0
\]
In other words, transferring $\varepsilon$ attention from $z_i$ to $z_h$ achieves a higher utility than the original allocation and hence we have reached a contradiction. An allocation such that $\sigma_z^2 < \sigma_h^2$ and $\sigma_z^2 < \sigma_j^2 < \sigma_h^2$ cannot be the solution to the maximization problem.

On the other hand, consider the case where the agent chooses an allocation $\sigma_j^2 \leq \sigma_z^2 < \sigma_h^2$. I will show that the mirror information acquisition strategy, i.e. $\{\sigma_j^2, \sigma_h^2\}$ achieves higher utility than $\{\sigma_z^2, \sigma_h^2\}$ and thus reach another contradiction. Consider the difference in utilities from these two information acquisition choices:

$$U(\sigma_h^2, \sigma_j^2) - U(\sigma_z^2, \sigma_h^2) = \int_{\ln(\sigma_z^2)}^{\ln(\sigma_h^2)} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_f^2, \exp(x))d(x) - \int_{\ln(\sigma_j^2)}^{\ln(\sigma_h^2)} \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(x), \sigma_h^2)d(x) < 0$$

The inequality follows for the same reasoning as in Case IV - $\frac{\partial U}{\partial \ln(\sigma_f^2)}(y, \exp(\ln(\sigma^2))), x < \frac{\partial U}{\partial \ln(\sigma_f^2)}(y, \exp(\ln(\sigma^2))), \forall x, y$, and thus we have reached another contradiction. Therefore, we have proven that it is never optimal to set $\sigma_j^2 < \sigma_z^2$ and $\sigma_h^2 > \sigma_z^2$. This means that an interior solution is optimal only if it sets $\sigma_h^2 \leq \sigma_z^2$. With this in mind, I will now show that in any interior solution it must be the case that $\sigma_h^2 < \sigma_j^2$. If an interior solution implies an allocation where $\sigma_j^2 > \sigma_z^2$ it is trivial that $\sigma_h^2 < \sigma_z^2$, since we just proved that any interior solution must be such that $\sigma_h^2 \leq \sigma_z^2$. So let us consider what kind of interior solutions can obtain if $\sigma_j^2 \leq \sigma_h^2$. Any interior solution requires

$$A \frac{\sigma_z^2}{(\sigma_h^2 + \sigma_z^2)^2} = B \frac{\sigma_j^2}{(\sigma_j^2 + \sigma_z^2)^2}$$

Remember that $A < B < 0$, while the other terms are positive, hence for the equality to hold we need

$$\frac{\sigma_h^2}{(\sigma_h^2 + \sigma_z^2)^2} < \frac{\sigma_j^2}{(\sigma_j^2 + \sigma_z^2)^2}$$

However, the function $\frac{\sigma^2}{(\sigma^2 + \sigma_{jz}^2)}$ is increasing in $\sigma^2$ for $\sigma^2 \leq \sigma_{jz}^2$ hence the above inequality can only hold if $\sigma_h^2 < \sigma_j^2$. Hence for any interior solution, it must be the case that $\sigma_h^2 < \sigma_j^2$. This is also trivially true for the corner solution where all attention is allocated to $z_h$ and hence this concludes the proof that at any solution, it is always the case that $\sigma_h^2 < \sigma_j^2$.

Lastly, we still need to prove that the solution of the maximization problem is unique. The proof amounts to showing that for any combination of parameters there is at most 1 possible interior solution and that this interior solution achieves the exact same utility as the corner solution only in the case where they correspond to the same allocations, and are thus the same solution.

Let us start by proving there is at most only one feasible interior solution. At any interior solution, it must be the case that

$$\frac{\partial U}{\partial \ln(\sigma_h^2)} - \frac{\partial U}{\partial \ln(\sigma_j^2)} = 0$$
Using the expressions for these derivatives, it follows that

\[ \frac{\partial U}{\partial \ln(\sigma_h^2)} - \frac{\partial U}{\partial \ln(\sigma_f^2)} = A \frac{\sigma_h^2}{(\sigma_h^2 + \sigma_z^2)^2} - B \frac{\sigma_f^2}{(\sigma_f^2 + \sigma_z^2)^2} \]

\[ = \frac{A\sigma_h^2(\sigma_f^2 + \sigma_z^2)^2 - B\sigma_f^2(\sigma_h^2 + \sigma_z^2)^2}{(\sigma_h^2 + \sigma_z^2)^2(\sigma_f^2 + \sigma_z^2)^2} \]

\[ = 0 \]

Which is true if and only if

\[ A\sigma_h^2(\sigma_f^2 + \sigma_z^2)^2 - B\sigma_f^2(\sigma_h^2 + \sigma_z^2)^2 = 0 \]

Now use the fact that the information capacity constraint is binding to get that \( \sigma_f^2 = \frac{\sigma_h^2}{\sigma_h^2 \exp(2\kappa)} \), substitute this expression on the left hand side above, expand and combine all terms to arrive at the following:

\[ (-B\sigma_z^4 + \exp(2\kappa)A\sigma_z^4)\sigma_h^4 + 2\sigma_z^2\sigma_e^2(A - B)\sigma_h^2 + \sigma_z^4(A\sigma_z^4 \exp(-2\kappa) - B\sigma_e^4) = 0 \]

Notice that the left hand side is a second order polynomial in \( \sigma_h^2 \), and we are interested in its strictly positive roots. We are only looking for positive solutions because \( \sigma_h^2 = 0 \) would require infinite information transfer and will violate the entropy constraint - such interior solutions are not feasible. The roots can be determined by standard root finding techniques, but a more intuitive and perhaps a bit less tedious approach is as follows. Let \( P(\sigma_h^2) \) be the quadratic polynomial of interest and notice that \(^{31}\)

\[ P(0) = \sigma_z^4(A\sigma_z^4 \exp(-2\kappa) - B\sigma_e^4) \geq 0, \quad \text{if} \quad \frac{\sigma_z^4}{\sigma_h^2 \exp(2\kappa)} \leq \frac{B}{A} \]

\[ < 0, \quad \text{if} \quad \frac{\sigma_z^4}{\sigma_h^2 \exp(2\kappa)} > \frac{B}{A} \]

\[ P'(0) = 2\sigma_z^4\sigma_e^2(A - B) < 0 \]

\[ P''(\sigma_h^2) = -B\sigma_z^4 + \exp(2\kappa)\sigma_e^4A \geq 0, \quad \text{if} \quad \frac{\sigma_z^4}{\sigma_h^2 \exp(2\kappa)} \geq \frac{A}{B} \]

\[ < 0, \quad \text{if} \quad \frac{\sigma_z^4}{\sigma_h^2 \exp(2\kappa)} < \frac{A}{B} \]

\(^{31}\)Of course while the chosen notation only makes it explicit that the polynomial is a function of \( \sigma_h^2 \), it is important to keep in mind that it depends on all other model parameters, e.g. \( \sigma_z^2, \kappa \), etc., as well.

47
First, let's handle the special case where the coefficient on \( \sigma^4_h \) is 0 and this is in fact a linear function. This happens when \( \frac{\sigma^4_h}{\sigma^4_z \exp(2\kappa)} = \frac{A}{B} > \frac{B}{A} \), hence \( P(0) < 0 \). And since the function is decreasing there is no \( \sigma^2_h > 0 \) such that \( P(\sigma^2_h) = 0 \). Second, consider \( \frac{\sigma^4_h}{\sigma^4_z \exp(2\kappa)} = \frac{B}{A} < \frac{A}{B} \). In this case, \( P(0) = 0 \), the polynomial is concave and decreasing at 0, hence there are no strictly positive solutions.

Now consider \( P''(\sigma^2_h) \neq 0 \). There are a few different cases to consider. First, let \( \frac{\sigma^4_z}{\sigma^4_h \exp(2\kappa)} > \frac{A}{B} > \frac{B}{A} \). This implies that \( P(0) < 0 \) and the polynomial is convex, therefore there is exactly one \( \sigma^2_h > 0 \) such that \( P(\sigma^2_h) = 0 \). Now assume instead that \( \frac{B}{A} < \frac{\sigma^4_h}{\sigma^4_z \exp(2\kappa)} < \frac{A}{B} \). In this case, \( P(0) < 0 \), the polynomial is concave and \( P'(0) < 0 \) all of which tell us that there exist no \( \sigma^2_h > 0 \) such that \( P(\sigma^2_h) = 0 \). Lastly, consider \( \frac{\sigma^4_z}{\sigma^4_h \exp(2\kappa)} < \frac{B}{A} \). In this case \( P(0) \geq 0 \) and the polynomial is concave and decreasing at 0, hence there is exactly one \( \sigma^2_h > 0 \) such that \( P(\sigma^2_h) = 0 \).

To summarize, we have the following:

\[
\begin{align*}
\frac{\sigma^4_z}{\sigma^4_h \exp(2\kappa)} &> \frac{A}{B} \quad \Rightarrow 1 \text{ positive solution} \\
\frac{\sigma^4_z}{\sigma^4_h \exp(2\kappa)} &\in \left[ \frac{B}{A}, \frac{A}{B} \right] \quad \Rightarrow \text{No positive solutions} \\
\frac{\sigma^4_z}{\sigma^4_h \exp(2\kappa)} &< \frac{B}{A} \quad \Rightarrow 1 \text{ positive solution}
\end{align*}
\]

Thus, for any combination of parameters we can have at most one feasible interior solution. Next, I shall that the global maximizer is always unique by comparing the feasible interior solution with the corner solution in the following three exhaustive cases.

Case 1: Let \( \frac{\partial U}{\partial \ln(\sigma^2_h)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right) = \frac{\partial U}{\partial \ln(\sigma^2_f)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right) \), or in other words, this is the situation where at the corner the derivative in terms of home information is equal to the derivative in terms of foreign information. Thus, this allocation also corresponds to the only interior solution, hence the corner and the interior solutions are one and the same in this case. The unique optimal solution in this case is the corner solution of allocating all attention to the home fundamental.

Case 2: Let \( \left| \frac{\partial U}{\partial \ln(\sigma^2_h)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right) \right| > \left| \frac{\partial U}{\partial \ln(\sigma^2_f)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right) \right| \). In this case, at the corner, the marginal benefit of an extra unit of home information is strictly higher than the marginal benefit of foreign information. Hence,

\[
P(\frac{\sigma^2_z}{\exp(2\kappa)}) = \frac{\partial U}{\partial \ln(\sigma^2_h)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right) - \frac{\partial U}{\partial \ln(\sigma^2_f)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right) < 0
\]

48
The polynomial whose roots characterize the interior solution is negative at the corner. Now I will show that the polynomial is also negative at \( \sigma_h^2 = \sigma_e^2 \) and thus because the polynomial has at most one zero for \( \sigma_h \geq 0 \) there cannot be any interior solutions where \( \sigma_h^2 \leq \sigma_e^2 \). This is enough to conclude that there are no optimal interior solutions, because it was already shown that the corner solution dominates all allocations where \( \sigma_h^2 > \sigma_e^2 \).

\[
P(\sigma_e^2) = (-B\sigma_e^4 + \exp(2\kappa)A\sigma_e^4)\sigma_e^4 + 2\sigma_e^4\sigma_e^4(A - B) + \sigma_e^4\left(\frac{A\sigma_e^4}{\exp(2\kappa)} - B\sigma_e^4\right)
= \sigma_e^4\sigma_e^4(-B + \frac{A\sigma_e^4\exp(2\kappa)}{\sigma_e^4} + 2(A - B) + A\frac{\sigma_e^4}{\exp(2\kappa)\sigma_e^4} - B)
= \sigma_e^4\sigma_e^4(2(A - B) + A\frac{\sigma_e^4\exp(2\kappa)}{\sigma_e^4} + \frac{\sigma_e^4}{\exp(2\kappa)\sigma_e^4}) - 2B)
\leq \sigma_e^4\sigma_e^4(2(A - B) + 2A - 2B)
= 4\sigma_e^4\sigma_e^4(A - B)
< 0
\]

The first inequality follows from the fact that \( y + \frac{1}{y} \geq 2 \) when \( y \geq 0 \) and the second inequality makes use of the fact that \( A < B < 0 \). Note that this result does not rely on any parameter restrictions - it is always the case that the derivative in respect to home information, evaluated at \( \sigma_h^2 = \sigma_e^2 \), is more negative than the derivative in respect to foreign information. Therefore, since \( P(\sigma_e^2) < 0 \), \( P(\frac{\sigma_2^2}{\exp(2\kappa)}) < 0 \) and there can be at most only one interior solution, it follows that there is no \( \sigma_2^2 \in \left[ \frac{\sigma_2^2}{\exp(2\kappa)}, \sigma_e^2 \right] \) such that \( P(\sigma_2^2) = 0 \) and hence there is no interior solution that could possibly be truly optimal. Thus, the corner solution is the unique maximum again.

Case 3: Lastly, consider the case when \( |\frac{\partial U}{\partial \ln(\sigma_h^2)}(\frac{\sigma_2^2}{\exp(2\kappa)}, \sigma_e^2)| < |\frac{\partial U}{\partial \ln(\sigma_f^2)}(\frac{\sigma_2^2}{\exp(2\kappa)}, \sigma_e^2)| \). In this situation, at the corner solution, the marginal benefit of home information is smaller than the marginal benefit of foreign information and hence

\[
P(\frac{\sigma_2^2}{\exp(2\kappa)}) = \frac{\partial U}{\partial \ln(\sigma_h^2)}(\frac{\sigma_2^2}{\exp(2\kappa)}, \sigma_e^2) - \frac{\partial U}{\partial \ln(\sigma_f^2)}(\frac{\sigma_2^2}{\exp(2\kappa)}, \sigma_e^2) > 0
\]

Since the derivatives are continuous functions, it follows that there exists a \( \bar{\varepsilon} > 0 \) such that for all \( \varepsilon \in (0, \bar{\varepsilon}) \):

\[
\frac{\partial U}{\partial \ln(\sigma_h^2)}(\exp(\sigma_2^2) - 2\kappa + \varepsilon, x) > \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(y, \ln(\sigma_2^2) - \varepsilon)), \forall x, y
\]

Therefore it will be beneficial to shift this \( \bar{\varepsilon} \) attention from \( z_h \) to \( z_f \). Formally, I consider going from \( \{ \frac{\sigma_2^2}{\exp(2\kappa)}, \sigma_e^2 \} \) to \( \{ \frac{\sigma_2^2(\bar{\varepsilon})}{\exp(2\kappa)}, \frac{\sigma_2^2}{\exp(\bar{\varepsilon})} \} \). 49
Proof. A.3 Proposition 3:
increasing returns whenever acquiring information about the forecastable fundamentals (both home and foreign) exhibits
of learnable uncertainty, otherwise, for
Proof. A.2 Proposition 2:
information than foreign information.
problem has a unique solution such that
as well.
Since the polynomial is continuous, there exists a
because by the assumption of Case 3,
is achieved by the interior solution. Notice that the interior solution is guaranteed to exist
Therefore, the corner solution is not optimal in this case and hence the unique maximum
is the solution to the information acquisition problem whenever
A<
\frac{\partial U}{\partial \ln(\sigma^2_h)}(x, \frac{\sigma^2_z}{e^x}) d\ln(x) - \int_{\frac{2}{e^x}}^{\sigma^2_z} \frac{\partial U}{\partial \ln(\sigma^2_h)}(e^{\frac{\sigma^2_z}{e^x}}, x) d\ln(x)
> 0

Therefore, the corner solution is not optimal in this case and hence the unique maximum
is achieved by the interior solution. Notice that the interior solution is guaranteed to exist
because by the assumption of Case 3, \( P(\frac{\sigma^2_z}{\exp(2x)}) > 0 \) and as derived previously \( P(\sigma^2_z) < 0 \).
Since the polynomial is continuous, there exists a \( \sigma^2 \in [\frac{\sigma^2_z}{\exp(2x)}, \sigma^2_z] \) for which \( P(\sigma^2) = 0 \) and
the interior solution is \( \sigma^2_h = \sigma^2 \). Again, this means that the maximum is unique in this case as well.
This concludes the proof of Proposition 1. It was shown that the information acquisition
problem has a unique solution such that \( \sigma^2_h < \sigma^2_f \), i.e. the agent always acquires more home
information than foreign information.

A.2 Proposition 2:
Proof. The second derivatives of the utility function \( U \) in terms of \( \ln(\sigma^2_h) \) and \( \ln(\sigma^2_f) \) are:
\[
\frac{\partial^2 U}{\partial \ln(\sigma^2_h)^2} = B \frac{\sigma^2_h (\sigma^2_e - \sigma^2_f)}{(\sigma^2_h + \sigma^2_e)^3}
\]
\[
\frac{\partial^2 U}{\partial \ln(\sigma^2_f)^2} = A \frac{\sigma^2_f (\sigma^2_e - \sigma^2_f)}{(\sigma^2_f + \sigma^2_e)^3}
\]
Since \( A < 0, B < 0 \) it follows that \( \frac{\partial^2 U}{\partial \ln(\sigma^2_i)^2} \geq 0 \) whenever \( \sigma^2_i \geq \sigma^2_e \), and \( \frac{\partial^2 U}{\partial \ln(\sigma^2_i)^2} < 0 \) otherwise, for \( i \in \{h, f\} \). Hence utility is convex in both signals' precisions whenever the size
of learnable uncertainty, \( \sigma^2_i \), is greater than the size of non-learnable uncertainty, \( \sigma^2_e \). Thus, acquiring information about the forecastable fundamentals (both home and foreign) exhibits
increasing returns whenever \( \sigma^2_i \geq \sigma^2_e \) and decreasing returns otherwise.

A.3 Proposition 3:
Proof. (i) As detailed in the Proof of Proposition 1, the corner allocation where \( \sigma^2_h = \frac{\sigma^2_z}{\exp(2x)} \)
is the solution to the information acquisition problem whenever \( \frac{\partial U}{\partial \ln(\sigma^2_h)}(\frac{\sigma^2_z}{\exp(2x)}, \sigma^2_z) \leq \frac{\partial U}{\partial \ln(\sigma^2_f)}(\frac{\sigma^2_z}{\exp(2x)}, \sigma^2_z) \). Using the analytical expression for the derivatives, this condition
amounts to
\[
A \frac{\sigma^2_z}{e^{2x} (\frac{\sigma^2_z}{e^x} + \sigma^2_e)^2} \leq B \frac{\sigma^2_z}{(\sigma^2_z + \sigma^2_e)^2}
\]
50
In particular, it is true that when $\kappa = 0$ the above inequality is strict. Moreover, as proved in Proposition 2, the function $f(x) = \frac{x}{(x+\sigma_2^2)}$ is increasing for $x \geq \sigma_2^2$ and decreasing otherwise. Hence, the above expression can hold with an equality for one and only one value of $\kappa$. Call this particular value of $\kappa$, $\kappa_i$ and notice that it is a function of all parameters of the model and in particular $\gamma$, $\delta$, $\sigma_2^2$ and thus we can write $\kappa_i(\gamma, \delta, \sigma_2^2)$ to make this dependence explicit. And since the left hand side of the above is always strictly less than the right hand side when $\kappa < \kappa_i(\gamma, \delta, \sigma_2^2)$ it follows from the proof of Proposition 1 that in this case the optimum is achieved at the corner solution $\sigma_h^2 = \frac{\sigma_f^2}{\exp(2\kappa_i)}$, $\sigma_f^2 = \sigma_2^2$. Thus, for $\kappa \leq \kappa_i(\gamma, \delta, \sigma_2^2)$:

$$
[\ln(\sigma_f^2) - \ln(\sigma_h^2)] = \frac{1}{2} [\ln(\sigma_f^2) - \ln(\sigma_h^2) + 2\kappa_i] = \kappa_i
$$

Clearly, $\Lambda$ is an increasing function of $\kappa$ for $\kappa \leq \kappa_i(\gamma, \delta, \sigma_2^2)$.

On the other hand, whenever $\kappa > \kappa_i(\gamma, \delta, \sigma_2^2)$ we have $\frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(2\kappa_i), \sigma_f^2) > \frac{\partial U}{\partial \ln(\sigma_f^2)}(\exp(2\kappa_i), \sigma_2^2)$ and the maximum is described by the unique interior allocation which makes the two partial derivatives equal and exhausts the whole information capacity. Let $\kappa \leq \kappa_1 < \kappa_2$ and consider the respective optimal allocations $\{\sigma_h^2, \sigma_f^2\}$ and $\{\sigma_h^2, \sigma_f^2\}$. I will show that the corresponding information asymmetry is smaller in the allocation corresponding to $\kappa_2$ as compared to the allocation under $\kappa_1$. There are two cases to analyze.

**Case 1:** Let $\sigma_f^2 \geq \sigma_h^2$. Recall that the interior solution $\sigma_h^2$ is the unique positive root of the following polynomial:

$$
(-B\sigma_z^4 + \exp(2\kappa_1)A\sigma_z^4)\sigma_h^4 + 2\sigma_z^2\sigma_h^2(A - B)\sigma_h^2 + \sigma_z^4(A\sigma_z^4 \exp(-2\kappa_1) - B\sigma_z^4)
$$

Call the polynomial $P_{\kappa}(\sigma_h^2)$ where this time the notation makes it explicit that the polynomial depends on the parameter $\kappa$. Given that the optimal solution for $\kappa_1$ is $\{\sigma_h^2, \sigma_f^2\}$, it must be the case that $P_{\kappa_1}(\sigma_h^2) = 0$. The proof will proceed by showing that $P_{\kappa_2}(\sigma_h^2) \geq 0$, which will be enough because from the proof of Proposition 1 we know that $P_{\kappa_2}(\sigma_h^2) < 0$.

Consider the derivative of the polynomial in terms of $\kappa$:

$$
\frac{\partial P_{\kappa_1}(\sigma_h^2)}{\partial \kappa} = 2e^{2\kappa_1}A\sigma_z^4\sigma_h^4 - 2A\sigma_z^8e^{-2\kappa_1}
$$

(2)

$$
= 2e^{2\kappa_1}A\sigma_z^4(\sigma_h^4 - \frac{\sigma_z^8}{e^{2\kappa_1}\sigma_z^2})
$$

(3)

From the binding entropy constraint, we can get the following relationship: $\sigma_f^2 = \frac{\sigma_z^4}{e^{2\kappa_1}\sigma_z^2}$. And since we have assumed that $\sigma_f^2 \geq \sigma_h^2$, it follows that $\frac{\sigma_z^4}{e^{2\kappa_1}\sigma_h^2} \geq \sigma_h^2$ and hence $\frac{\sigma_z^4}{e^{2\kappa_1}\sigma_h^2} \geq \sigma_h^2$. 51
Substituting this in equation (3) above, we obtain that
\[
\frac{\partial P_{\kappa_1}(\sigma_h^2)}{\partial \kappa} = 2e^{2\kappa} A\sigma_e^4(\sigma_h^4 - \frac{\sigma_e^8}{e^{2\kappa}\sigma_e^8}) \geq 0
\]
and thus, \(P_{\kappa_1}(\sigma_h^2) \geq 0\). From the proof of Proposition 1 we also know that \(P(\sigma_e^2) < 0\) and therefore, the unique \(\sigma_h^2\) such that \(P_{\kappa_2}(\sigma_h^2) = 0\) is in the interval \([\sigma_h^2, \sigma_e^2]\). Moreover, since the information acquisition constraint must be binding, it is the case that \(\sigma_h^2 \leq \sigma_f^2\) (otherwise the constraint will be lax), so we have that
\[
\sigma_h^2 \leq \sigma_h^2 \\
\sigma_f^2 \geq \sigma_f^2
\]
and only one of the inequalities can hold with an equality. Therefore, we can conclude that
\[
\frac{1}{2}(\ln(\sigma_f^2) - \ln(\sigma_h^2)) > \frac{1}{2}(\ln(\sigma_f^2) - \ln(\sigma_h^2))
\]
which proves that \(\Lambda\) is a decreasing function of \(\kappa\) for \(\kappa \geq \bar{\kappa}\).

Case 2: Let \(\sigma_f^2 < \sigma_e^2\). Under this condition, it is immediate to see that
\[
\frac{\partial P_{\kappa_1}(\sigma_h^2)}{\partial \kappa} = 2e^{2\kappa} A\sigma_e^4(\sigma_h^4 - \frac{\sigma_e^8}{e^{2\kappa}\sigma_e^8}) < 0
\]
which implies that \(P_{\kappa_2}(\sigma_h^2) < 0\). On the other hand, from the proof of Proposition 1 we know that there exists a unique interior solution and that \(P_{\kappa_2}(\sigma_e^2) < 0\) as well, which lets us conclude that the solution \(\sigma_h^2 < \sigma_e^2\). Having determined that, we can use Proposition 2 to show that it must also be the case that \(\sigma_f^2 < \sigma_f^2\), or otherwise the partial derivatives would not be equal, which is a condition of the interior solution. Hence, we have shown that whenever \(\sigma_f^2 < \sigma_e^2\), increasing \(\kappa\) leads to increasing both home and foreign information (i.e., decreasing \(\sigma_h^2\) and \(\sigma_f^2\) both). This means that we can rewrite \(\{\sigma_h^2, \sigma_f^2\}\) as
\[
\sigma_h^2 = \sigma_h^2 e^{-\nu_h} \\
\sigma_f^2 = \sigma_f^2 e^{-\nu_f}
\]
where \(\nu_h + \nu_f = \kappa_2 - \kappa_1\) and \(\nu_h, \nu_f > 0\). One can think of the \(\nu\)'s as the extra home and foreign information, respectively, the agent will acquire if we were to increase his capacity from \(\kappa_1\) to \(\kappa_2\). I will show that \(\nu_h < \nu_f\) by the way of contradiction, so to this end assume the opposite: \(\nu_h \geq \nu_f\). Since both \(\{\sigma_h^2, \sigma_f^2\}\) and \(\{\sigma_h^2, \sigma_f^2\}\) characterize interior solutions, it is the case that

52
Then, we have

\[ \frac{\partial U}{\partial \ln(\sigma_h^2)}(\sigma_h^2, \sigma_f^2) = \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, \sigma_f^2) \]

\[ \frac{\partial U}{\partial \ln(\sigma_h^2)}(\sigma_h^2, \sigma_f^2) = \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, \sigma_f^2) \]

Subtract the two lines from one another, move all terms on the left side and re-arrange:

\[ \left( \frac{\partial U}{\partial \ln(\sigma_h^2)}(\sigma_h^2, \sigma_f^2) - \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, \sigma_f^2) \right) - \left( \frac{\partial U}{\partial \ln(\sigma_h^2)}(\sigma_h^2, \sigma_f^2) - \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_h^2, \sigma_f^2) \right) = \]

\[ = \int_{x_0}^{x_1} \frac{\partial^2 U}{\partial \ln(\sigma_h^2)^2}(x, \sigma_f^2) \, dx - \int_{x_0}^{x_1} \frac{\partial^2 U}{\partial \ln(\sigma_f^2)^2}(x, \sigma_h^2) \, dx = \]

The first equality follows from \( \frac{\partial U}{\partial \ln(\sigma_f^2)}(\sigma_f^2, x) = \frac{\partial U}{\partial \ln(\sigma_h^2)}(\sigma_h^2, y) \) for all \( x, y > 0 \) as the cross partials are 0.

Computing the second derivative expressions yields:

\[ \frac{\partial^2 U}{\partial \ln(\sigma_h^2)^2}(\sigma_h^2, \sigma_f^2) = A \frac{\sigma_h^2 - \sigma_f^2}{(\sigma_h^2 + \sigma_f^2)^3} \]

\[ \frac{\partial^2 U}{\partial \ln(\sigma_f^2)^2}(\sigma_h^2, \sigma_f^2) = B \frac{\sigma_f^2 - \sigma_h^2}{(\sigma_f^2 + \sigma_h^2)^3} \]

and since \( \sigma_h^2 < \sigma_f^2 < \sigma_e^2 \) it follows that

\[ \frac{\partial^2 U}{\partial \ln(\sigma_h^2)^2}(\sigma_h^2, \sigma_f^2) < \frac{\partial^2 U}{\partial \ln(\sigma_f^2)^2}(\sigma_h^2, \sigma_f^2) \]

Then, we have
As which is a contradiction. Thus, we must have \( B \) solution hence sequence of optimal information allocations. When sequence of real numbers that diverges to infinity and generality assume that it is one of the variances converges to a number strictly greater than 0, and without loss of to zero - we just need to prove that both do. To do so, assume to the contrary that constraint always holds, it must be the case that at least one of the two variances goes of \( \nu_f \), decreases with \( \kappa \).

(ii) This follows directly from the Proof of Proposition 1. For \( \kappa \leq \bar{\kappa} \) the solution is characterized by the corner allocation where \( \sigma_h^2 = \frac{\sigma_f^2}{\exp(2\nu_0)} \) and \( \sigma_f^2 = \sigma_z^2 \). In the case of \( \kappa > \bar{\kappa} \) we have an interior solution, and from the Proof of Proposition 1 we know that this only happens when \( \frac{\partial U}{\partial \ln(\sigma_h^2)} \left( \frac{\sigma_h^2}{\exp(2\nu_0)}, \sigma_f^2 \right) = \frac{\partial U}{\partial \ln(\sigma_f^2)} \left( \frac{\sigma_h^2}{\exp(2\nu_0)}, \sigma_f^2 \right) \) and \( \kappa \) is lower than \( \bar{\kappa} \). Since the information acquisition constraint is always binding, this implies that \( \sigma_f^2 < \sigma_z^2 \) and we can conclude that for \( \kappa > \bar{\kappa} \) the agent pays a positive amount of attention to foreign information.

(iii) First, I will establish that \( \sigma_h^2 \to 0, \sigma_f^2 \to 0 \) as \( \kappa \to \infty \). Since the information acquisition constraint always holds, it must be the case that at least one of the two variances goes to zero - we just need to prove that both do. To do so, assume to the contrary that one of the variances converges to a number strictly greater than 0, and without loss of generality assume that it is \( \sigma_f^2 \to \sigma^2 > 0 \). In order to formalize things, let \( \{\kappa_n\} \) be a sequence of real numbers that diverges to infinity and \( \{\sigma_{h,n}^2, \sigma_{f,n}^2\} \) be the corresponding sequence of optimal information allocations. When \( \kappa_n \) is big, we obtain the interior solution hence

\[
\frac{\sigma_{h,n}^2}{(\sigma_{h,n}^2 + \sigma_f^2)^2} = \frac{\sigma_{f,n}^2}{(\sigma_{f,n}^2 + \sigma_z^2)^2}
\]

As \( n \to \infty \), the left hand side converges to 0, while the right hand side converges to \( B^2 \sigma^2_f < 0 \). Hence there exists an integer \( N \) such that for all \( n \geq N \),
which means that the sequence \( \{ \sigma^2_{h,n}, \sigma^2_{f,n} \} \) does not characterize solutions to the information problem and we have reached a contradiction - it must be the case that both variances converge to 0. Having established this, consider again the fact that the two partial derivatives must be equal:

\[
A \frac{\sigma^2_{h,n}}{(\sigma^2_{h,n} + \sigma^2_e)^2} = B \frac{\sigma^2_{f,n}}{(\sigma^2_{f,n} + \sigma^2_e)^2}
\]

\[
\Rightarrow A \frac{(\sigma^2_{f,n} + \sigma^2_e)^2}{B (\sigma^2_{h,n} + \sigma^2_e)^2} = \frac{\sigma^2_{f,n}}{\sigma^2_{h,n}}
\]

As \( n \to \infty \) the left hand side converges to \( \frac{A}{B} \) hence we get that

\[
\frac{\sigma^2_{f,n}}{\sigma^2_{h,n}} \to \frac{A}{B}
\]

And since natural log is a continuous function we get that

\[
\Lambda = \frac{1}{2} (\ln(\sigma^2_f) - \ln(\sigma^2_h)) \to \frac{1}{2} \ln(\frac{A}{B})
\]

A.4 Proposition 4:

Proof. By the proof of Proposition 3, we know that \( \bar{\kappa} \) is the value of \( \kappa \) such that

\[
\frac{\partial U}{\partial \ln(\sigma^2_h)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right) = \frac{\partial U}{\partial \ln(\sigma^2_f)} \left( \frac{\sigma^2_z}{\exp(2\kappa)}, \sigma^2_z \right)
\]

which reduces to

\[
A \frac{\sigma^2_z}{e^{2\kappa}(\sigma^2_h + \sigma^2_z)^2} = B \frac{\sigma^2_z}{\sigma^2_z + \sigma^2_e}
\]  

(4)

First, lets analyze the effect of \( \delta \) on the value of \( \bar{\kappa} \). The expression for \( A \) and \( B \) are given in the proof of Proposition 1, and we can see from them that only \( A \) is a function of \( \delta \). In particular, we can show that \( A \) is a decreasing function of \( \delta \), hence if we have \( \delta_1 < \delta_2 \) it follows that \( A(\delta_1) > A(\delta_2) \). Let \( \bar{\kappa}_1 \) and \( \bar{\kappa}_2 \) be the values of \( \kappa \) which make (4) hold with an equality, when \( A \) is calculated under \( \delta_1 \) and \( \delta_2 \) respectively. The right hand side will be the same under both \( \delta_1 \) and \( \delta_2 \), but since \( A(\delta_1) > A(\delta_2) \), it follows that

55
well and roots are given by the measure of (eventually disappearing) increasing returns to information acquisition. If and only if $\sigma_2^2 < \sigma_2^e$, the solution is always interior and the agent always acquires both home and foreign information for strictly concave, the solution is always interior and the agent always acquires both home and foreign information. Notice that the fraction of the total variance of the financial assets which is forecastable, i.e. 

\[ \bar{\sigma}_x^2 \text{ and } \bar{\sigma}_y^2, \] 

it must be the case that $y < x$. Therefore, since

\[ \frac{x}{(x + \sigma_2^2)^2} = \frac{y}{(y + \sigma_2^2)^2}, x < \sigma_2^2, y < \sigma_2^2 \]

it must be the case that $\sigma_2^2 > \sigma_2^e$ which means that $\bar{\kappa}_1 < \bar{\kappa}_2$ and thus we conclude that $\bar{\kappa}(\delta, \gamma, \alpha)$ is increasing in $\delta$.

Next I analyze the relationship between $\bar{\kappa}$ and $\gamma$. $A$ is a decreasing function of $\gamma$ as well and $B$ is still not affected by changes in $\gamma$, hence using the exact same arguments as above, we can conclude that $\bar{\kappa}(\gamma, \delta, \alpha)$ is increasing in $\gamma$.

Lastly, let us consider how changes in $\alpha$ affect the value of $\bar{\kappa}$. By definition $\alpha$ is the fraction of the total variance of the financial assets which is forecastable, i.e. $\sigma_2^2 = \alpha \sigma_2^y$. First, notice that $\bar{\kappa}$ exists only when $\alpha > \frac{1}{2}$, as otherwise the information acquisition problem is strictly concave, the solution is always interior and the agent always acquires both home and foreign information. A situation in which the agent acquires only home information and ignores foreign information for $\kappa \leq \bar{\kappa}$ can only occur if $\alpha > \frac{1}{2}$ because this would imply a measure of (eventually disappearing) increasing returns to information acquisition.

Next, rewrite (4) in terms of $\alpha$ by using the expressions $\bar{\sigma}_x^2 = \alpha \sigma_2^y$ and $\bar{\sigma}_y^2 = (1 - \alpha) \sigma_2^y$: 

\[ A = \frac{\sigma_2^2}{(\bar{\sigma}_x^2 + (1 - \alpha) e^\kappa)^2} = B \]

Expanding the expression in parentheses and combining terms I get:

\[ (1 - \alpha)e^{2\kappa} - \left( \frac{A}{B} \right)^\frac{1}{2} e^\kappa + \alpha = 0 \]

The left-hand side of the above expression is a second order polynomial in $\exp(\kappa)$ which roots are given by

\[ e^\kappa = \left( \frac{A}{B} \right)^\frac{1}{2} \pm \frac{\left( \frac{A}{B} - 4\alpha(1 - \alpha) \right)^\frac{1}{2}}{2(1 - \alpha)} \]

Since $A < B$, we know that \( \frac{\partial U}{\partial m(\sigma_2^2) (\sigma_2^2, \sigma_2^y)} < \frac{\partial U}{\partial m(\sigma_2^y) (\sigma_2^2, \sigma_2^y)} \). Moreover, by Proposition 2...
\[
\frac{\partial U}{\partial \ln(\sigma_0^2)}(\sigma_0^2, \sigma_z^2) \text{ is at first a decreasing and then an increasing function of } \kappa \geq 0. \text{ Therefore, there can be only one } \kappa \geq 0 \text{ such that (4) is satisfied. With this in mind, it is immediate that the appropriate root of the above polynomial is then}
\]
\[
e^\kappa = \frac{\left(\frac{A}{B}\right)^{\frac{1}{2}} + \left(\frac{A}{B} - 4\alpha(1-\alpha)\right)^{\frac{1}{2}}}{2(1-\alpha)}
\]

This is the case because the other root is strictly lower than this one, hence if the other one was to be obtained at \(\kappa \geq 0\) then there would be two positive solutions to equation (4) which will be a contradiction. Thus the above equation gives an implicit function for \(\tilde{\kappa}\), and its partial derivative in respect to \(\alpha\) will help determine the relationship between \(\alpha\) and \(\tilde{\kappa}\):

\[
\frac{\partial e^\kappa}{\partial \alpha} = \frac{\left(\frac{\partial |A|}{\partial \alpha} + \frac{\partial |A|}{\partial \alpha} - 4(1-\alpha) + 4\alpha \right)2(1-\alpha) + 2\left(\frac{A}{B}\right)^{\frac{1}{2}} + \left(\frac{A}{B} - 4\alpha(1-\alpha)\right)^{\frac{1}{2}}}{4(1-\alpha)^2}
\]

The sign of the above is determined by the sign of its numerator:

\[
\left(\frac{\partial |A|}{\partial \alpha} + \frac{\partial |A|}{\partial \alpha} - 4(1-\alpha) + 4\alpha \right)2(1-\alpha) + 2\left(\frac{A}{B}\right)^{\frac{1}{2}} + \left(\frac{A}{B} - 4\alpha(1-\alpha)\right)^{\frac{1}{2}} =
\]

\[
\frac{A - B}{(|AB|^2)^{\frac{1}{2}}} + \frac{A - B}{(|AB|^2)^{\frac{1}{2}}} - 4(1-2\alpha)(1-\alpha) + 2\left(\frac{A}{B}\right)^{\frac{1}{2}} + 2\left(\frac{A}{B} - 4\alpha(1-\alpha)\right)^{\frac{1}{2}} >
\]

\[
-(\frac{A}{B})^{\frac{1}{2}} + 2\left(\frac{A}{B}\right)^{\frac{1}{2}} + \frac{2\frac{A}{B} - \frac{A}{B} + 1 - 4 + 12\alpha - 8\alpha^2 - 8\alpha + 8\alpha^2}{\left(\frac{A}{B} - 4\alpha(1-\alpha)\right)^{\frac{1}{2}}} >
\]

\[
\frac{2(2\alpha - 1)}{(\frac{A}{B} - 4\alpha(1-\alpha))^{\frac{1}{2}}} > 0
\]

The first equality follows from

\[
\frac{\partial |A|}{\alpha}(1-\alpha) = -\gamma(1-\delta)(\mu - pR)\sigma_z^2 - \gamma^2(1-\delta)^2\sigma_z^2
\]

\[= A - B
\]

the first inequality follows from the fact that \(-B > 0\), the second from \(\frac{A}{B} > 1\) and the last inequality is due to \(\alpha > \frac{1}{2}\). These derivations let us conclude that

\[
\frac{\partial e^{\kappa}}{\partial \alpha} > 0
\]

which implies \(\tilde{\kappa}\) is increasing in \(\alpha\).

Lastly, let’s analyze how \(\Lambda\) behaves given changes in \(\gamma\) or \(\delta\) - imagine again we move
from \(\gamma_1\) to \(\gamma_2\) or from \(\delta_1\) to \(\delta_2\), where \(\gamma_1 < \gamma_2\) and \(\delta_1 < \delta_2\). Let \(\kappa_1\) correspond to \(\gamma_1\) (or \(\delta_1\)) and \(\kappa_1\) correspond to \(\gamma_2\) (or \(\delta_2\)). Then, if \(\kappa < \kappa_1\), \(\Lambda = \kappa\) and there is no change when moving from \(\gamma_1\) to \(\gamma_2\) (or from \(\delta_1\) to \(\delta_2\)). When \(\kappa_1 \leq \kappa \leq \kappa_2\), then clearly \(\Lambda\) increases as under the new parameter values \(\Lambda = \kappa\), but \(\Lambda < \kappa\) under \(\gamma_1\) or \(\delta_1\). Lastly, consider \(\kappa > \kappa_2\). In this case, the solution must be the unique positive root of the polynomial \(P(\sigma^2_z)\) defined in the Proof of Proposition 1. Notice that

\[
\frac{\partial P(\sigma^2_h)}{\partial \gamma} = \frac{\partial A}{\partial \gamma} (e^{2\kappa} \sigma^4_h \sigma^4_z + 2\sigma^2_z \sigma^2_h \sigma^2_z + \sigma^6_z e^{-2\kappa}) < 0
\]

and

\[
\frac{\partial P(\sigma^2_f)}{\partial \delta} = \frac{\partial A}{\partial \delta} (e^{2\kappa} \sigma^4_h \sigma^4_z + 2\sigma^2_z \sigma^2_h \sigma^2_z + \sigma^6_z e^{-2\kappa}) < 0
\]

where the inequalities follow from \(\frac{\partial A}{\partial \gamma} < 0\) and \(\frac{\partial A}{\partial \delta} < 0\). The above two results, combined with \(P(\sigma^2_z) < 0\) and the fact that the unique positive solution must be strictly less than \(\sigma^2_z\) imply that the solution \(\sigma^2_{k_2}\), under \(\gamma_2\) (or respectively \(\delta_2\)), must be strictly lower than \(\sigma^2_h\), the solution under \(\gamma_1\) (or respectively \(\delta_1\)). Therefore, we conclude that \(\Lambda\) is increasing in \(\gamma\) and \(\delta\) for any given \(\kappa\).

**A.5 Proposition 5**

**Proof.** First, I will address the case \(\sigma^2_z = 0\). Under this condition, the objective function becomes:

\[
U(\sigma^2_h, \sigma^2_f) = \frac{1}{2} (-2 + \frac{\sigma^2_z}{\sigma^2_h} (1 + \frac{(\mu - pR)^2}{\sigma^2_z}) + \frac{\sigma^2_z}{\sigma^2_f} (1 + \frac{(\mu - pR)^2}{\sigma^2_z}))
\]

and the information capacity constraint is:

\[
\frac{1}{2} (\ln(\text{Var}(z_h|s)) - \ln(\sigma^2_h) + \ln(\sigma^2_z) - \ln(\sigma^2_f)) \leq \kappa
\]

All of the notation is the same as the above, except for \(\sigma^2_h\) which is now defined as \(\sigma^2_h = \text{Var}(z_h|s, \eta_h)\) - the posterior variance of \(z_h\) after observing both the exogenous, free signal \(s\) and the endogenously chosen signal \(\eta_h\). The change in the information capacity constraint comes from the assumption that the agents receive the exogenous home signal \(s\) for free. All uncertainty reduction delivered by \(s\) does not count against the information capacity that constrains the choice of the endogenous signals \(\eta_h\) and \(\eta_f\). The agent is restrained in the amount of uncertainty he can reduce, over and above the uncertainty reduction delivered by the exogenous home signal \(s\).

Notice that the constraint can also be rewritten as:

\[
\frac{\text{Var}(z_h|s)\sigma^2_z}{\sigma^2_h \sigma^2_f} \leq \exp(2\kappa)
\]

Hence the agent’s problem is to maximize a sum, under a product constraint and the no forgetting constraints. The solution to this type of problem (as also derived by
Van Nieuwerburgh and Veldkamp (2010)) is the corner solution where the agent spends his whole constraint on the term with the highest linear weight. A quick inspection of the objective function shows that the linear weight on the home signal term \( \frac{\text{Var}(z_h|s)}{\sigma^2_h} \) is:

\[
\frac{\sigma^2_z}{\text{Var}(z_h|s)}(1 + \frac{(\mu - pR)^2}{\sigma^2_z})
\]

and the linear weight on the foreign signal \( \frac{\sigma^2_z}{\sigma^2_f} \) is:

\[
(1 + \frac{(\mu - pR)^2}{\sigma^2_z})
\]

Because of the exogenous information advantage over the home asset, we have that \( \frac{\sigma^2_z}{\text{Var}(z_h|s)} > 1 \) and the linear weight on the home signal is bigger. The unique solution of this problem is the corner solution where the agent allocates all of his information capacity to the home asset and \( \kappa = \kappa \). This is the same result as in Van Nieuwerburgh and Veldkamp (2010).

On the other hand, if \( \sigma^2_e > 0 \) the agent solves the following problem:

\[
\max_{\sigma^2_h, \sigma^2_f} U(\sigma^2_h, \sigma^2_f) = \frac{1}{2}(-2 + \frac{\sigma^2_z + \sigma^2_e}{\sigma^2_h + \sigma^2_e}(1 + \frac{(\mu - pR)^2}{\sigma^2_z + \sigma^2_e}) + \frac{\sigma^2_z + \sigma^2_e}{\sigma^2_f + \sigma^2_e}(1 + \frac{(\mu - pR)^2}{\sigma^2_z + \sigma^2_e}))
\]

s.t.

\[
\frac{1}{2}(\ln(\text{Var}(z_h|s)) - \ln(\sigma^2_h) + \ln(\sigma^2_e) - \ln(\sigma^2_f)) \leq \kappa
\]

\[-\infty \leq \ln(\sigma^2_h) \leq \ln(\sigma^2_e), \quad -\infty \leq \ln(\sigma^2_f) \leq \ln(\sigma^2_e)
\]

This problem does not share the same convexity properties as in the case of \( \sigma^2_e = 0 \) and the corner solution is not always optimal. In particular, I will show that for \( \kappa < \kappa \) the problem obtains the corner solution, for \( \kappa = \kappa \) the agent is indifferent between any attention allocation (multiplicity of solutions), and for \( \kappa > \kappa \) the interior solution is the unique solution to the problem. Where the \( \kappa \) is defined as

\[
\kappa = \frac{1}{2}(\ln(\text{Var}(z_h|s)) - \ln(\sigma^2_e) + \ln(\sigma^2_e) - \ln(\sigma^2_e))
\]

Intuitively, this is just the information capacity which allows the agent to set both the home and foreign asset posterior variances equal to \( \sigma^2_e \) (the size of unlearnable uncertainty). To prove the statements above, start by analyzing the Karush-Kuhn-Tucker conditions, which imply that there are three possible optimal allocations - an interior solution at which the partial derivatives are equal, and two corner solutions - one where the agent allocates all of the capacity to the home signal, and the opposite one where all of the capacity is allocated to the foreign signal. Consider first \( \kappa < \kappa \) and note that for such \( \kappa \) it is impossible to set both \( \sigma^2_h < \sigma^2_e \) and \( \sigma^2_f < \sigma^2_e \). On the other hand, an interior solution must satisfy

\[
\frac{\sigma^2_h}{(\sigma^2_h + \sigma^2_e)^2} = \frac{\sigma^2_f}{(\sigma^2_f + \sigma^2_e)^2}
\]
and by the proof of Proposition 2 both the LHS and the RHS are decreasing whenever \( \sigma_h^2 \geq \sigma_e^2 \) or respectively, \( \sigma_f^2 \geq \sigma_e^2 \). Without loss of generality, assume that \( \sigma_f^2 \geq \sigma_e^2 \). Then, it is straightforward to show (using the same arguments as in the Proof of Proposition 1) that a small deviation from any interior point to \( \left\{ \sigma_h^2, \sigma_e^2 \right\} \) would achieve a strictly higher utility. Intuitively, this is because in this situation the marginal benefit of increasing the precision of any of the signals is increasing, and with a convex function like this the interior solution cannot be optimal. Thus, the only thing we need to determine is which of the two corner solutions achieves higher utility:

\[
U \left( \frac{\text{Var}(z_h|s)}{e^{2\kappa}}, \sigma_e^2 \right) - U \left( \sigma_h^2, \frac{\sigma_e^2}{e^{2\kappa}} \right) =
\]

\[
= (\sigma_h^2 + \sigma_e^2) \left( 1 + \frac{(\mu - pR)^2}{\text{Var}(z_h|s)} \right) \left( \frac{1}{\sigma_h^2 + \sigma_e^2} + \frac{1}{\sigma_h^2 + \sigma_e^2} - \frac{1}{\text{Var}(z_h|s) + \sigma_e^2} - \frac{1}{\sigma_h^2 + \sigma_e^2} \right)
\]

\[
= C(\sigma_h^2 - \text{Var}(z_h|s))(\sigma_h^2 + \sigma_e^2)(\text{Var}(z_h|s) + \sigma_e^2) - \frac{1}{(\sigma_h^2 + \sigma_e^2)(\text{Var}(z_h|s) + \sigma_e^2)}
\]

where \( C = (\sigma_h^2 + \sigma_e^2)(1 + \frac{(\mu - pR)^2}{\sigma_h^2 + \sigma_e^2}) \). To determine the sign of the last term in parentheses, notice that

\[
e^{2\kappa}(\sigma_h^2 + \sigma_e^2)(\text{Var}(z_h|s) + \sigma_e^2) - (\text{Var}(z_h|s) + e^{2\kappa}\sigma_e^2)(\sigma_h^2 + e^{2\kappa}\sigma_e^2) = (e^{2\kappa} - 1)(\text{Var}(z_h|s)\sigma_h^2 - e^{2\kappa}\sigma_e^4)
\]

\[
> (e^{2\kappa} - 1)(\text{Var}(z_h|s)\sigma_h^2 - e^{2\kappa}\sigma_e^4)
\]

\[
= 0
\]

where the inequality follows from \( \kappa < \bar{\kappa} \), and the last equality from \( e^{2\kappa} = \frac{\text{Var}(z_h|s)\sigma_e^2}{\sigma_h^2} \).

The last result lets us conclude that \( U \left( \frac{\text{Var}(z_h|s)}{e^{2\kappa}}, \sigma_e^2 \right) - U \left( \sigma_h^2, \frac{\sigma_e^2}{e^{2\kappa}} \right) > 0 \) and hence the corner solution where the agent allocates all of his information capacity to learning about the home asset is optimal whenever \( \kappa < \bar{\kappa} \). Consequently, for \( \kappa < \bar{\kappa} \) we have \( \Lambda = \kappa \).

Now consider \( \kappa = \bar{\kappa} \). First, using the argument of the previous paragraph we see that the two mirror corner solutions achieve the same level of utility. On the other hand, using the fact that the information constraint is always binding (\( \sigma_f^2 = \frac{\text{Var}(z_h|s)\sigma_e^2}{\sigma_h^2 e^{2\kappa}} \)) we can derive that any interior solution is a positive root of the following quadratic equation (derivations are same as in the Proof of Proposition 1):

\[
P(\sigma_h^2) = \sigma_h^4 \left( - \text{Var}(z_h|s)\sigma_e^2 + e^{2\kappa}\sigma_e^4 \right) + \left( \frac{\text{Var}(z_h|s)^2\sigma_e^4}{e^{2\kappa}} - \text{Var}(z_h|s)\sigma_e^2 \right)
\]

But at \( \kappa = \bar{\kappa} \) the RHS of the above reduces to 0, hence any \( \sigma_h^2 \) is an interior solution. This, plus the fact that the partial derivatives are equal at all interior solutions is enough to conclude that the agent is indifferent between any information allocation.
Lastly, consider $\kappa > \kappa$. First, note that $\frac{\sigma^2}{(\sigma^2 + \sigma_z^2)^2}$ is increasing in $\sigma^2$ and thus decreasing in $\kappa$. Therefore, it is straightforward to show that $|\frac{\partial U}{\partial \ln(\sigma^2)}(\text{Var}(z_h|s), \sigma^2_h)| < |\frac{\partial U}{\partial \ln(\sigma^2)}(\text{Var}(z_h|s), \sigma^2_z)|$ and that $|\frac{\partial U}{\partial \ln(\sigma^2)}(\text{Var}(z_h|s), \sigma^2_z)| > |\frac{\partial U}{\partial \ln(\sigma^2)}(\text{Var}(z_h|s), \sigma^2_x)|$, which intuitively means that at each corner solution the marginal benefit of acquiring information for the other fundamental is bigger than the marginal benefit of pushing the corner even further out. Then, by a similar argument as in the Proof of Proposition 1, it follows that the corner solutions are dominated by the interior solution. Moreover, the interior solution is unique, again by an argument in the same spirit as the ones presented in the Proof of Proposition 1, and it is given by the condition $\sigma^2_h = \sigma^2_j$.

Thus, we have $\frac{\sigma^2_h}{\sigma^2_j} = 1$, and

$$
\Lambda = \frac{1}{2}[(\ln(\text{Var}(z_h|s))) - (\ln(\sigma^2_h))) - (\ln(\sigma^2_j)) - (\ln(\sigma^2_j))] < 0
$$

And finally, consider how $\tilde{\kappa} = \frac{\text{Var}(z_h|s)\sigma^2}{\sigma^2_z}$ varies with $\sigma^2_\alpha$ (the variance of the error in the exogenous signal $s$) and $\alpha$, where $\alpha$ is defined as above - $\sigma^2_z = \alpha \sigma^2_y$. Using the standard expression for posterior variance $\text{Var}(z_h|s) = \frac{\sigma^2_\alpha \sigma^2_y}{\sigma^2_\alpha + \sigma^2_y}$ I obtain:

$$
\frac{\partial \tilde{\kappa}}{\partial \sigma^2_\alpha} = \frac{\sigma^2_\alpha \sigma^4_y}{((\sigma^2_\alpha + \sigma^2_z)^4)} > 0
$$

and

$$
\frac{\partial \tilde{\kappa}}{\partial \alpha} = \frac{\alpha^2(1 - \alpha)^2 \sigma^2_\alpha \sigma^2_y + 2\alpha(1 - \alpha)^2 \sigma^4_y + 2(1 - \alpha)\sigma^2_\alpha \sigma^2_y}{((1 - \alpha)^2(\alpha \sigma^2_y + \sigma^2_\alpha)^2)} > 0
$$

Thus, $\tilde{\kappa}$ is increasing with $\sigma^2_\alpha$ and $\alpha$. \hfill \Box

B Appendix B: Model Robustness

B.1 Thinking About Mutual Funds

An information asymmetry explanation of portfolio concentration rests on the assumption that the individual investors trade only on differing private information sets. However, one could wonder why there is no mutual funds industry, presumably with a higher ability to acquire information than the individual investors, that would offer well diversified investment opportunities to the investors. After all, information asymmetry is more plausible at the household level, rather than at the fund level, because the fund managers have a lot more time and resources to spend on acquiring all relevant information. This is an important criticism on the literature explaining the home bias through information asymmetry as a whole, and in this section I outline a few different reasons for why introducing a mutual funds
industry would not invalidate the results of this paper. The key intuition is that as long as the objective function or the quality of the different fund managers are not observable the investors will prefer funds that specialize and ex-ante commit to focusing either on “domestic” or “foreign” securities. Moreover, the investors will then rely on their private information sets to allocate their money between the different specialized funds, and this would again result in biased portfolios.

This result is in line with the data. Coeurdacier and Rey (2013) show that the great majority of US Funds either have 90 – 100% of their assets invested domestically or 0 – 10%. The distribution of US funds is remarkably bi-modal, with two pronounced mass points at the two ends and almost no funds in between. This suggests that specialization is a defining feature of the US mutual funds industry and the resulting landscape basically leaves investors with a binary choice of a “home” and a “foreign” investment. Coeurdacier and Rey (2013) also document that there is a considerably amount of heterogeneity in the fund industry across countries, but the general finding that funds tend to cluster around “domestic-only” and “foreign-only” types holds true for most countries. A careful analysis of the structure of the fund industry could perhaps help shed more light on the issue of home bias, but is beyond the scope of this Appendix. Here I will just show that specialization of funds arises naturally because the final decision of what to do with one’s money is the investor’s, and hence his private information set is never irrelevant.

An obvious issue with the introduction of mutual funds into the model is that the investors would also face a principal-agent problem. This arises because the investors cannot be sure that the objective of the fund managers coincides with their own. For example, the fund manager might just be running a Ponzi scheme and be stealing from the invested money. Or less abrasively, the fund manager might just be using a different stochastic discount factor than the investor and thus may choose an inefficient portfolio, from the view point of a single investor, despite his superior information and good intentions. In short, investing in mutual funds adds an additional layer of uncertainty from the view point of the individual investor because the fund manager’s portfolio choice is also viewed as stochastic.

In the extreme case where investors do not trust the managers to make their investment decisions at all, there will only be any demand for index funds with pre-announced portfolio allocations (e.g. a fund that mirrors the S&P 500). In this case, it will be optimal for all funds to fully specialize in “domestic only” and “foreign only”, as this will attract the most customers (I am assuming funds try to maximize the number of their customers). One way to combat this mistrust is for the mutual funds to announce particular investment mandates in their prospectuses. For example, a fund might decide to invest “at least x% of equity in domestic stocks”. A commitment to a particular type of investment strategy like this makes the fund manager’s portfolio choice less uncertain to the investors, and leaves a smaller scope for a principal-agent problem. And in fact mandates like this are quite common in the mutual funds industry (see Coeurdacier and Rey (2013)) - most funds do have a clearly defined investment “type”.

However, mandates that restrict the fund managers’ portfolio choices also diminishes their ability to use their superior information to earn better returns than the individual investors. The individual investors are still not able to just turn their money over to a better informed professional, because the professional now faces extra constraints which hinder his performance. Thus, the investor’s private information set is not irrelevant. In fact, the agent
will use his own beliefs to choose whether it is best to invest in funds specializing in home or foreign securities. Hence, the overall decision of whether to invest at home or abroad is still tied to the individual investor’s information set, which is biased towards home information.

Apart from the principal-agent problem, another issue arises with the introduction of mutual funds if investors are not able to observe the quality of the different managers. In a similar setting of information asymmetry, Dziuda and Mondria (2012) show that investors will prefer funds that specialize in domestic securities, because they will be better able to judge a manager’s quality. So even though the managers themselves are equally well informed about both home and foreign securities, they choose to focus on the domestic market because this attracts more customers. Hence, this is a framework in which the information asymmetry plaguing the individual investors generates home bias even in the presence of sophisticated, well informed professional fund managers.

At the end of the day, although interesting, a detailed analysis of the mutual funds industry and its implications for the home bias is beyond the scope of this paper. This sections only purports to sketch a few considerations which could leave investors reliant on their private information sets and thus holding biased portfolios, even in the presence of well informed fund managers. And while this section only considered the issue from a theoretical point, it is important to note that in the data mutual funds do adhere to published investment mandates and tend to specialize into broadly defined ”domestic” and ”foreign” funds. Thus, for one reason or another, the equilibrium we observe in practice does not seem to offer an easily identifiable “optimal” fund to investors. And in such a case individual biases and information frictions are certain to play a role.

C Appendix C: Empirical Results Appendix

C.1 The list of countries in the dataset

The countries covered by the dataset are: Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Israel, Italy, Japan, Korea, Latvia, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

C.2 Data Description

All variables are on an annual basis and have been collected for the time period 2001-2008. The measure of home bias I use is the previously defined EHB index, which is standard in the literature. I follow standard practice and calculate the EHB index for each country-year pair combining portfolio data from the IMF’s Coordinated Portfolio Investment Survey (CPIS) with domestic stock market capitalization data from the World Bank.

The data on Labor Income comes from the OECD database. I use the series Total Labor Cost, which is Total Employee Compensation adjusted for self-employment income and the amount of self-employed people in the economy. The Financial Wealth proxies are
constructed from CPIS and World Bank data. The Total Equity Holdings measure comes
directly from CPIS and World Bank data series and is constructed as follows:

\[
\text{Total Equity Holdings} = \text{Domestic Portfolio Equity Assets} + \text{Foreign Portfolio Equity Assets}
\]

where Domestic Portfolio Equity Assets is computed as is standard in the literature:

\[
\text{Domestic Portfolio Equity Assets} = \text{Domestic Market Capitalization} - \text{Foreign Portfolio Equity Liabilities}
\]

The data on Domestic Market Capitalization is from the World Bank, and data on Foreign Portfolio Equity Assets and Liabilities is from the CPIS.

Constructing Total Financial Assets takes into consideration the holdings of portfolio
debt securities as well, but is a little harder because there is no data on the total Domestic
Market Capitalization of Debt Securities. Thus, I cannot compute a measure of Domestic
Portfolio Debt Assets from directly observable data. Instead, I assume that the share of the
total market value of domestic debt securities that is held by foreigners is the same as the
share of the stock market that is held by foreigners. With that assumption in hand, we can
express the total market value of domestic debt securities as:

\[
\text{Market Value of Domestic Debt Securities} = \frac{\text{Foreign Portfolio Debt Liabilities}}{1 - \text{DomesticShare}}
\]

Then, we can compute the Domestic Portfolio Debt Assets using the imputed value of the
Market Value of Domestic Debt Securities, and finally arrive at Total Financial Assets:

\[
\text{Total Financial Assets} = \text{Equity Holdings} + \text{Foreign Portfolio Debt Assets} + \text{Domestic Portfolio Debt Assets}
\]

Finally, I turn these variable into per person measures by dividing by the number
of adults, i.e. people over the age of 15, which data I obtain from the World Bank. The
Appendix shows that the results are virtually unchanged if instead I divide by total population.
However, the number of adults is theoretically more appealing because these are the people
that earn income and make financial decisions.

The Chinn-Ito index of financial openness is taken directly from Chinn and Ito (2006).
It is the first principal component of the four binary measures of capital controls that are
tracked by the IMF in their Annual Report on Exchange Arrangements and Exchange
Restrictions. As a robustness check, the Appendix also presents results when instead I use
the Lane-Milesi-Ferretti Index of Financial Openness, which is the sum of foreign assets and
liabilities over total GDP. The results are unchanged. Lastly, the real GDP per capita data
comes from the World Bank.

All income variables are reported in constant Purchasing Power Parity terms (PPP).
The results are virtually unchanged if instead the income variables are measured in constant
US dollars and a Table with estimates of these regressions is included in this Appendix.
C.3 Robustness Checks

C.3.1 Alternative Proxies for \( \kappa \)

Table 3: Alternative Proxies for \( \kappa \)

<table>
<thead>
<tr>
<th></th>
<th>Cell Phones</th>
<th>Computers (WDI)</th>
<th>Computers (OECD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log ( \kappa ) proxy</td>
<td>-0.088 (0.064)</td>
<td>0.001 (0.033)</td>
<td>-0.103 (0.212)</td>
</tr>
<tr>
<td>Log Labor Income</td>
<td>0.142 (0.093)</td>
<td>-0.225 (0.276)</td>
<td>-0.241 (0.466)</td>
</tr>
<tr>
<td>Log Total Fin Assets</td>
<td>-0.016 (0.018)</td>
<td>-0.021 (0.025)</td>
<td>0.021 (0.068)</td>
</tr>
<tr>
<td>Log Equity Assets</td>
<td>-0.001 (0.031)</td>
<td>0.059 (0.055)</td>
<td>0.034 (0.070)</td>
</tr>
<tr>
<td>Chinn-Ito Index</td>
<td>-0.053*** (0.018)</td>
<td>0.021 (0.035)</td>
<td>-0.057 (0.069)</td>
</tr>
<tr>
<td>Log RGDP 2001 p.c.</td>
<td>-0.286*** (0.104)</td>
<td>-0.064 (0.321)</td>
<td>-0.043 (0.305)</td>
</tr>
<tr>
<td>( N )</td>
<td>35</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.657</td>
<td>0.609</td>
<td>0.594</td>
</tr>
</tbody>
</table>

OLS point estimates with heteroskedasticity robust standard errors in parentheses. ***,** and * denote significance at the 1%, 5% and 10% respectively.

Table 3 presents regression results for three alternative proxies of information capacity \( \kappa \) - number of cell phones per 100 people, the number of computers per 100 people (from the World Bank’s WDI database) and the percentage of households with a PC (OECD Database). The coefficient on cell phones is negative but is not significant (significant at the 9% level in a one-sided test) and none of the coefficient estimates for the PC variables is found to be anywhere close to being statistically significant. Hence, the estimation results based on these proxies are not as strong as the ones found when using Internet users per 100 people. But perhaps this is not too surprising as the number of cell phones per capita is at best a problematic measure of information processing capacity. The Internet’s main use and greatest benefit is fast and convenient procurement of information, while in the period 2001-2008 this was a function of only secondary or even tertiary importance for a cell phone (remember that the iPhone was only released in 2007 for the first time). Nevertheless, I am including these results for completeness due to the dearth of reliable proxies for information technology.

On the other hand, the regression results with the two computer based proxies appear to suffer from multicollinearity problems. This is mainly because both computers data series have rather poor coverage and thus these regressions are only based on 24 and 19 observations respectively. Moreover, the WDI series was discontinued in 2003 and the 3rd and 4th columns are estimated with data from 2001 to 2003, while the OECD series starts in 2006 and thus includes the years 2006 to 2008. The sharp reduction in available observations leads to a multicollinearity caused breakdown in the estimation of the last four columns of Table 3. All four regressions find highly significant \( R^2 \) in the neighborhood of 0.6 but only one significant coefficient among all of them, which is a textbook sign of multicollinearity and unfortunately
renders all individual coefficient estimates unreliable. Perhaps in the future the OECD will expand the breadth of their computers data series coverage and future researchers could reexamine the question, but the currently available computer series do not yield themselves to a reliable analysis.

C.3.2 Robustness Checks on the Population, Income and Financial Openness Measures

Table 4: Controls Robustness Checks

<table>
<thead>
<tr>
<th></th>
<th>Per Capita</th>
<th>Constant 2000 USD</th>
<th>Lane-Milesi-Ferretti Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Internet Users</td>
<td>-0.122**</td>
<td>-0.118**</td>
<td>-0.122**</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.052)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Log Labor Income</td>
<td>0.258***</td>
<td>0.222*</td>
<td>0.309***</td>
</tr>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.117)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Log Total Fin Assets</td>
<td>-0.024</td>
<td>-0.026</td>
<td>0.021*</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.018)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Log Equity Assets</td>
<td>0.011</td>
<td>-0.014</td>
<td>0.073**</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.044)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Financial Openness Index</td>
<td>-0.068***</td>
<td>-0.059***</td>
<td>-0.064***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.017)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Log RGDP_{2001 p.c.}</td>
<td>-0.307***</td>
<td>-0.351***</td>
<td>-0.320***</td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.085)</td>
<td>(0.106)</td>
</tr>
</tbody>
</table>

OLS point estimates with heteroskedasticity robust standard errors in parentheses.
***, ** and * denote significance at the 1%, 5% and 10% respectively.

Table 4 provides regression results when income variables are standardized by total population, instead of number of adults (columns (1) and (2)), when all monetary variables are expressed in 2000 US dollars (columns (3) and (4)) and when the Lane-Milesi-Ferretti Index of financial openness is used instead of the Chinn-Ito index. The results are robust to these alternative specifications. The coefficient on Internet users per capita is negative, statistically significant and of the same magnitude as the specifications reported in the main body of the paper.

On the other hand, in columns (1) through (4) the coefficient of Labor Income is positive, significant and of the same magnitude as in the main body of the paper and the coefficient on financial wealth is again insignificant. Using the Lane-Milesi-Ferretti Index of financial openness produces slightly different results. In that case the coefficient on labor income is insignificant, but the coefficient on Financial Assets is found to be positive and marginally significant in column 5 and significant at the 5% level in column 6. It appears that in this case the estimated effect of the hedging motive dominates the estimated effect of the information acquisition motive, as the relative size of labor income seems to have a negative relationship with the home bias. Thus, under the Lane-Milesi-Ferretti Financial Openness index the second empirical restriction is not rejected and hence the evidence of
relative labor income is consistent with both models. Nevertheless, we can still reject the first restriction as the coefficient on Internet users is found to be negative and the regressions still rule in favor of the endogenous information asymmetry model.

### C.3.3 Regressing on the Ratio of Labor and Financial Incomes

Table 5: Restricted Labor Income and Financial Holdings Coefficient

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Internet Users</td>
<td>-0.236***</td>
<td>-0.219***</td>
<td>-0.229***</td>
<td>-0.132***</td>
<td>-0.161***</td>
<td>-0.078*</td>
<td>-0.077*</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.035)</td>
<td>(0.031)</td>
<td>(0.042)</td>
<td>(0.037)</td>
<td>(0.048)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Log Labor Income</td>
<td>0.036</td>
<td>0.044**</td>
<td>0.024</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Assets</td>
<td>(0.029)</td>
<td>(0.019)</td>
<td>(0.018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor Share</td>
<td>-0.110</td>
<td>0.089</td>
<td>0.323</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.377)</td>
<td>(0.274)</td>
<td>(0.218)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinn-Ito Index</td>
<td>-0.076***</td>
<td>-0.073***</td>
<td>-0.056**</td>
<td>-0.0506***</td>
<td>-0.118***</td>
<td>-0.165***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.017)</td>
<td>(0.040)</td>
<td>(0.035)</td>
<td></td>
</tr>
<tr>
<td>Log RGDP$_{2001}$ p.c.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.475</td>
<td>0.503</td>
<td>0.513</td>
<td>0.628</td>
<td>0.61</td>
<td>0.661</td>
<td>0.653</td>
</tr>
</tbody>
</table>

OLS point estimates with heteroskedasticity robust standard errors in parentheses. ***, ** and * denote significance at the 1%, 5% and 10% respectively.

Table 5 presents two new sets of results. First, it includes estimates when I impose the restriction $\beta_l = \beta_f$, i.e. use the log of the ratio of Labor Income over Financial Wealth as a single regressor, and second it presents results when I use the standard measure of labor income share of GDP. In both cases there are no significant changes to the results. The coefficient on log Internet users is found to be negative, significant and of the same magnitude as before, and the coefficient on the income composition ratio is generally found to be positive but insignificant (although it is significant at the 5% level in Column (4)). Again, the results are virtually the same as in the main body of the paper and support the conclusions reached there.

### C.3.4 Including A Measure of Non-Tradable Consumption

Table 6 presents regression results when I also control for the relative size of non-tradables in consumption expenditures. This regressor is motivated by the theoretical work of Tesar (1993) and I follow the empirical work of Lewis (1996) and Pesenti and Van Wincoop (2002) when I construct the empirical measure of non-tradable consumption. The data I use is from the OECD and consists of aggregate consumption expenditures broken down into different types such as Food, Health Services, Housing Expenses etc. Identifying non-tradable consumption in the data is tricky and no one approach is perfect, hence the table contains results for two different proxies of non-tradable consumption. The first two columns present results when I use the disaggregated data to arrive at a figure for non-tradable consumption and then I
Table 6: Results when controlling for non-tradable consumption

<table>
<thead>
<tr>
<th></th>
<th>Computing Non-Tradables</th>
<th>Computing Tradables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Internet Users</td>
<td>-0.179***</td>
<td>-0.173***</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Log Labor Income</td>
<td>0.394*</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>(0.224)</td>
<td>(0.287)</td>
</tr>
<tr>
<td>Log Total Fin Assets</td>
<td>0.003</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Log Non-tradable Con / Tradable Con</td>
<td>0.106</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.215)</td>
</tr>
<tr>
<td>Chinn-Ito Index</td>
<td>-0.060***</td>
<td>-0.059***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Log RGDP_{2001} p.c.</td>
<td>-0.627***</td>
<td>-0.601***</td>
</tr>
<tr>
<td></td>
<td>(0.202)</td>
<td>(0.199)</td>
</tr>
<tr>
<td>N</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.755</td>
<td>0.764</td>
</tr>
</tbody>
</table>

OLS point estimates with heteroskedasticity robust standard errors in parentheses.

***, ** and * denote significance at the 1%, 5% and 10% respectively.

subtract this from total consumption to obtain tradable consumption. The second pair of columns does the opposite - I use the disaggregated data to compute tradable consumption and then impute the non-tradable part. I follow Lewis (1996) and Pesenti and Van Wincoop (2002) as closely as possible in determining whether each type of disaggregated consumption is tradable or not. For example, I categorize Food, Beverages, Furniture and Vehicle purchases as tradable consumption and Health care, Financial and Restaurant services as non-tradable and so on.\(^{32}\)

C.3.5 Using a proxy for \( \kappa \) in the spirit of Mondria and Wu (2010)

Mondria and Wu (2010) also study the empirical relationship between \( \kappa \) and the Equity Home Bias but use a slightly different proxy for \( \kappa \). Instead of using a proxy for information technology level normalized by total population (like Internet Users per 1000 ppl) they further normalize this figure by real GDP per capita. Hence, they use Proxy for \( \kappa = \frac{\text{Information Tech per 1000 ppl}}{\text{Real GDP p.c.}} \) which they argue captures the average amount of information capacity per $1000 of economic activity. This extra normalization is appealing because the theoretical models studied here and in Mondria and Wu (2010) analyze only financial decisions and not the full array of decisions an individual must make in the real world. Thus, \( \kappa \) measures the information capacity that the agent allocates to financial decisions, and

\(^{32}\)The two approaches yield different results because the disaggregated consumption categories do not sum up to total consumption expenditures. The OECD only collects disaggregated data on certain types of consumption expenditures.
not necessarily his whole ability to process information. In light of this, an ideal empirical model would also attempt to control for how much of the total information capacity of the agent is allocated to financial decisions. Mondria and Wu (2010) attempt to do this by dividing by real GDP per capita, which they argue is a good proxy for how many other economic activities the agent needs to pay attention to. However, normalizing by real GDP per capita is not likely to capture the desired effect because it also proxies for a number of other important considerations. For example, real GDP per capita is an additional proxy for information capacity itself because a simple measure like Internet Users per capita cannot possibly characterize the whole information technology infrastructure in a given country. Furthermore, richer agents with large financial portfolios would likely prefer to apportion more (an not less) of their fixed information capacity to financial decisions as they have more at stake.

In general, it is not clear why Information Technology per capita would be a worse proxy for $\kappa$ than Information technology per $\$1000$ of real GDP. On the other hand, Information Technology per capita is an intuitive and readily available measure and this paper opts to include it in its raw form rather than apply any transformations to it. Instead, I include initial period real GDP per capita as a separate regressor to control for the initial differences between countries. This leads to an empirical framework that is a bit more flexible, and still controls for all important effects.

In the sake of completeness, however, Tables 7 and 8 present estimation results when the proxy for $\kappa$ is defined as in Mondria and Wu (2010) - $\frac{\text{Internet Users per 100 ppl}}{\text{Real GDP p.c.}}$. Table 7 uses all 35 countries in my data set, while Table 8 has only the 19 countries used by Mondria and Wu (2010). The conclusions of the main body of the paper are unchanged - the relationship between the home bias and information capacity is found to be negative and statistically significant. In fact, the size of the coefficient $\beta_\kappa$ is estimated to be about 4-5 times bigger (in absolute value) than the one found in the specifications in the main body of the paper. The only case in which I find a positive and significant $\beta_\kappa$ is when I regress Home Bias only on the proxy for $\kappa$ and the Chinn-Ito Financial Openness index. However, if I control for initial conditions (or add measures of labor and financial income) the sign flips back to being negative.

This is an interesting finding because Mondria and Wu (2010) find the opposite - a positive and significant $\beta_\kappa$. There are three possibilities for the discrepancy in the results. First, the datasets and data periods are quite different. Mondria and Wu (2010) use portfolio data from the IMF’s Balance of Payments Statistics and International Financial Statistics (IFS) databases over the span of 1988-2004, while I use data from the IMF’s Coordinated Portfolio Investment Survey for the years 2001-2008. There is only a little overlap in the available years of data. And the data itself is significantly different. The portfolio data in the IFS and the BPS is imputed based on reported capital flows between countries and is not based on direct observations of actual portfolio holdings. The CPIS, on the other hand, was commissioned by the IMF to specifically provide high quality portfolio data based on direct observations of portfolio holdings. In principle, it does not suffer from a lot of statistical measurement errors that could plague the IFS and BPS imputed numbers and is now the principal source of International Portfolio data for the literature (e.g. Coeurdacier and Rey (2013),Sercu and Vanpée (2008)). Thus, the different time span of the data and the different
Table 7: Results with Mondria and Wu (2010) normalization for $\kappa$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Internet Users</td>
<td>0.108***</td>
<td>-0.570**</td>
<td>-0.039</td>
<td>0.006</td>
<td>-0.560**</td>
<td>-0.5533**</td>
</tr>
<tr>
<td>Real GDP p.c.</td>
<td>(0.0223)</td>
<td>(0.287)</td>
<td>(0.164)</td>
<td>(0.164)</td>
<td>(0.284)</td>
<td>(0.271)</td>
</tr>
<tr>
<td>Log Labor Income</td>
<td>-0.062</td>
<td>-0.064</td>
<td>0.181</td>
<td>0.183</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
<td>(0.127)</td>
<td>(0.130)</td>
<td>(0.149)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Total Fin Assets</td>
<td>-0.039*</td>
<td>-0.0300</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Equity Assets</td>
<td>-0.012</td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.044)</td>
<td></td>
<td></td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>Chinn-Ito Index</td>
<td>-0.072***</td>
<td>-0.061***</td>
<td>-0.078***</td>
<td>-0.071***</td>
<td>-0.070***</td>
<td>-0.063***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.022)</td>
<td>(0.019)</td>
<td>(0.018)</td>
<td>(0.021)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Log RGDP 2001 p.c.</td>
<td>-0.563**</td>
<td></td>
<td></td>
<td>-0.689***</td>
<td>-0.736***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td></td>
<td></td>
<td>(0.253)</td>
<td>(0.240)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.570</td>
<td>0.645</td>
<td>0.602</td>
<td>0.575</td>
<td>0.676</td>
<td>0.659</td>
</tr>
</tbody>
</table>

OLS point estimates with heteroskedasticity robust standard errors in parentheses. ***,** and * denote significance at the 1%, 5% and 10% respectively.

Secondly, Mondria and Wu (2010) do not use labor and financial income data, nor control for initial GDP per capita in their specifications. Table 7 and 8 show that such considerations are important, as controlling for these effects has important implications for the sign and significance of $\beta_k$. Thirdly, Mondria and Wu (2010) run their regressions on a panel data set of 17 years and 19 countries, while I first time average the data and then use cross-sectional regressions. Theoretically, the cross-sectional regressions are the correct specification for testing the models considered in both this paper and Mondria and Wu (2010) because both are static models. At this point, the literature has not considered a dynamic model of endogenous information acquisition and portfolio choice and it is not clear what are the testable implications of such models. In any case, when I re-estimate Tables 7 and 8 on the the panel dimensions of my dataset I find once again that $\beta_k < 0$ and statistically significant.

Lastly, the standard OLS based fixed effects estimator that relies on cross-sectional variation for identification is known to suffer potentially significant biases in panels where the number of time periods available is of the same order as the available cross-sections and some of the variables are near unit root (like the Home bias for example). Perhaps an econometric issue could also help explain some part of the differences in the reported estimates.\footnote{On the other hand, the panel regressions I consider on the 2001-2008 dataset at hand are less likely to suffer from such problems for two reasons. First, I use GMM estimators that do not have asymptotic biases and secondly, the time periods (8) is much smaller than the number of cross sections (35).}
Table 8: Mondria and Wu (2010) normalization for $\kappa$ (19 countries sample)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Internet Users</td>
<td>-0.013</td>
<td>-1.39$^*$</td>
<td>-1.30$^*$</td>
<td>-1.45$^*$</td>
<td>-1.54$^*$</td>
<td>-1.81$^{**}$</td>
</tr>
<tr>
<td>Real GDP p.c.</td>
<td>(0.112)</td>
<td>(0.806)</td>
<td>(0.711)</td>
<td>(0.821)</td>
<td>(0.850)</td>
<td>(0.916)</td>
</tr>
<tr>
<td>Log Labor Income</td>
<td>-0.912</td>
<td>-1.23$^*$</td>
<td>-0.644</td>
<td>-0.891</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.578)</td>
<td>(0.655)</td>
<td>(0.523)</td>
<td>(0.644)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Total Fin Assets</td>
<td>-0.074$^*$</td>
<td>-0.073$^*$</td>
<td>(0.039)</td>
<td>(0.038)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Equity Assets</td>
<td>0.058</td>
<td>0.070</td>
<td>(0.062)</td>
<td>(0.066)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinn-Ito Index</td>
<td>-0.086$^*$</td>
<td>-0.086$^*$</td>
<td>-0.210$^{***}$</td>
<td>-0.118$^*$</td>
<td>-0.197$^{***}$</td>
<td>-0.100</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.045)</td>
<td>(0.051)</td>
<td>(0.068)</td>
<td>(0.062)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Log RGDP$_{2001}$ p.c.</td>
<td>-1.24$^*$</td>
<td>-0.517</td>
<td>-0.733</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.710)</td>
<td>(0.595)</td>
<td>(0.584)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.095</td>
<td>0.253</td>
<td>0.364</td>
<td>0.308</td>
<td>0.379</td>
<td>0.337</td>
</tr>
</tbody>
</table>

OLS point estimates with heteroskedasticity robust standard errors in parentheses.

$^{***}$, $^{**}$ and $^*$ denote significance at the 1%, 5% and 10% respectively.

### C.3.6 Evidence for a Non-Monotonic Relationship Between the Home Bias and $\kappa$

Proposition 3 implies that the home bias should have a non-monotonic relationship with $\kappa$, which is increasing for small $\kappa$ and decreasing for all $\kappa$ larger than a tipping point $\bar{\kappa}$. However, all regression specifications reported up to this point allow only for a monotonic (either positive or negative) relationship. The primary reason is the relatively small sample size (35 observations). Allowing for a more flexible, possibly non-monotonic, relationship with $\kappa$ would put a higher identification burden on the regression specification and one could run into problems of overfitting the available data. On the other hand, finding a negative $\beta_\kappa$ in the main specification which only allows for a linear effect is enough to conclude that the overall relationship between $\kappa$ and home bias is negative. And an overall negative relationship is enough to conclude that the data is at odds with the implications of a standard exogenous information advantage model. Moreover, a priori it seems unlikely that any significant portion of OECD countries in the 21st century would be informationally constrained, i.e. have $\kappa < \bar{\kappa}$. Thus, for the purposes of differentiating the exogenous information advantage model from the endogenous information asymmetry model, the main specification reported in the paper appears to be enough.

For the sake of robustness, however, it is interesting to also examine the possibility of a non-monotonic relationship. An attempt in this direction is given by Table 9 which reports results for a regression where I also include a squared Internet Users per 100 ppl term. The estimates show that there is indeed evidence of nonmonotonicity - home bias is increasing in $\kappa$ for low values and decreasing for high values. For example, the point estimates from the main specification (column 6) imply that the home bias is increasing in Internet Users if a country has less than roughly 24 Internet users per 100 ppl and decreasing for any values
above that. For the sample at hand, this implies that Brazil, Bulgaria, Mexico, Romania, and Turkey are the five OECD countries which fall into the information constrained set where $\kappa < \tilde{\kappa}$.

A visually appealing way to examine the potential non-monotonicity is to graphically look at the relationship between $\kappa$ and the home bias after controlling for other effects. In particular, I fit a 3rd order polynomial of Log Internet Users per 100 people on the residual from a regression of the Home bias on all other variables. The results are presented in Figure 7. The left panel shows the residual of the following regression

$$EHB_i = \text{const} + \beta_l \ln(\text{LabIncome}) + \beta_f \ln(\text{FinWealth}) + \beta'X_i + \varepsilon_i$$

plotted against Log Internet Users per 100 people and also draws the 3rd order polynomial fitted line. The second plot estimates the main regression specification of the model:

$$EHB_i = \text{const} + \beta_k \ln(InetU_i) + \beta_l \ln(\text{LabIncome}) + \beta_f \ln(\text{FinWealth}) + \beta'X_i + \varepsilon_i$$

and then looks at the residual Home Bias after I control for all variables except for the Internet Users per 100 people:

$$\varepsilon_r = EHB_i - \text{const} - \beta_l \ln(\text{LabIncome}) - \beta_f \ln(\text{FinWealth}) - \beta'X_i$$

The results from both exercises are pretty similar. The 3rd order polynomial has a well pronounced hump, where it is increasing for low values of Internet Users and decreasing for high values. In particular, in both figures the polynomial is increasing for the same five

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Internet Users</td>
<td>-0.018</td>
<td>0.194</td>
<td>0.420</td>
<td>0.705*</td>
<td>0.816*</td>
<td>0.831***</td>
<td>1.094***</td>
</tr>
<tr>
<td>(Log Internet Users)$^2$</td>
<td>-0.031</td>
<td>-0.042</td>
<td>-0.072</td>
<td>-0.111*</td>
<td>-0.126*</td>
<td>-0.131***</td>
<td>-0.167***</td>
</tr>
<tr>
<td>Log Labor Income</td>
<td>-0.139**</td>
<td>-0.234**</td>
<td>-0.036</td>
<td>-0.113</td>
<td>0.283***</td>
<td>0.212**</td>
<td></td>
</tr>
<tr>
<td>Log Total Fin Assets</td>
<td>-0.017</td>
<td>-0.032*</td>
<td>0.026</td>
<td>0.019</td>
<td>0.025</td>
<td>0.005</td>
<td>0.042</td>
</tr>
<tr>
<td>Log Equity Assets</td>
<td>-0.012</td>
<td>0.005</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.020)</td>
<td>(0.024)</td>
<td></td>
</tr>
<tr>
<td>Chinn-Ito Index</td>
<td>-0.077***</td>
<td>-0.0689***</td>
<td>-0.074***</td>
<td>-0.067***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log RGDP$^2001$ p.c.</td>
<td>-0.359***</td>
<td>-0.421***</td>
<td>(0.086)</td>
<td>(0.099)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N      | 35  | 35  | 35  | 35  | 35  | 35  | 35  |
$R^2$  | 0.4765 | 0.5719 | 0.5692 | 0.661 | 0.642 | 0.715 | 0.720 |

OLS point estimates with heteroskedasticity robust standard errors in parentheses.

***, ** and * denote significance at the 1%, 5% and 10% respectively.
countries found above (Brazil, Bulgaria, Mexico, Romania, and Turkey), and decreasing for all others.

Figure 7: Home Bias and $\kappa$