The Review of Financial Studies

Beyond Home Bias: International Portfolio Holdings and Information Heterogeneity

Filippo De Marco
Bocconi University, CEPR

Marco Macchiavelli
Federal Reserve Board

Rosen Valchev
Boston College

We show that international portfolios reflect the underlying heterogeneity in investors' beliefs. Using data on the foreign sovereign debt holdings of European banks matched with their forecasts on future bond yields, we find that expecting higher returns and having more accurate forecasts are associated with larger bond holdings. Crucially, the elasticity of portfolio holdings to expected returns is increasing in the precision of the forecast, implying that investors optimally exploit comparative advantages in information production. We rationalize the results in a model in which partial information specialization arises endogenously by introducing a degree of unlearnable uncertainty about asset payoffs. (*JEL* F21, F36, G11, D82)

Received December 2, 2020; editorial decision August 26, 2021 by Editor Stefano Giglio. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

The lack of international diversification in investors' portfolios, as first noted by French and Poterba (1991), is a prominent feature of the data. Most of the existing studies examining this issue have focused on the so-called "home bias," which refers to the surprisingly large portfolio share devoted to domestic assets. However, international portfolios also exhibit patterns of "foreign" biases, in

We are grateful to Philippe Bacchetta, Wilko Bolt, Stephen Cecchetti, Nicola Gennaioli, Tarek Hassan, Sebnem Kalemli-Ozcan, Nicola Limodio, Thomas Maurer, Alvaro Pedraza, Felipe Saffie, David Schumacher, Andrea Vedolin, Laura Veldkamp, and Francis Warnock and conference participants at AFFI 2017, ASSA 2019, Boston Macro Juniors 2019, HEC-McGill Winter Finance Workshop 2019, EFA 2018, FutFinInfo 2021, MFA 2018, Minnesota Finance Juniors Conference 2018, NBER IFM Summer Institute 2018, and SED 2019 for helpful comments. Jamie Grasing provided excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Board of Governors or the Federal Reserve System. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to valchev@bc.edu.

The Review of Financial Studies 35 (2022) 4387–4422

doi: 10.1093/rfs/hhab135 Advance Access publication December 11, 2021

the sense that investors under- and overweight individual foreign assets relative to one another (Chan et al. 2005). As a result, portfolios seem underdiversified not only because of the home bias effect but also because the share of wealth allocated to foreign investments is mainly concentrated in a few assets. Thus, while the classic home bias is an important feature of the data, it is not a sufficient statistic for the composition of international portfolios, and it is thus crucial to also confront international portfolio theories with data on the heterogeneity among individual foreign investments.

This paper takes a step in this direction by constructing and analyzing a new micro-level data set that combines data on both portfolio holdings of individual foreign assets and investors' beliefs about those specific assets' return characteristics. To do so, we compile a panel data set of European banks' portfolios of foreign sovereign debt, matching individual banks' holdings of long-term sovereign bonds to the corresponding 3-month-ahead forecasts of 10-year sovereign yields reported by the same banks in the Consensus Economics survey.[1] As bond yields are inversely proportional to bond prices, the yield forecasts proxy (reciprocally) for expected quarterly bond returns. Importantly, the panel nature of our data allow us to quantify the objective precision of individual forecasts by computing the average squared forecast error for each bank-country pair. As a result, we obtain proxies for both the first and second moments of investor beliefs about asset returns.

We find that both higher expected returns and greater forecast precision lead to larger portfolio holdings. Crucially, there is a positive interaction effect: the sensitivity of holdings to expected returns is greater when the underlying forecast is objectively more precise. This suggests that investors internalize and exploit their respective comparative advantages in information production, confirming the key prediction of information-based models of international portfolio allocations (e.g., Van Nieuwerburgh and Veldkamp 2009).

Furthermore, our findings can also help to pin down the specific structure of the information production technology, a crucial degree of freedom in the theoretical literature. In particular, the heterogeneity that we document in the precision of forecasts across different foreign assets is inconsistent with the typical information structure in the literature, which implies that agents fully specialize in producing additional information only about the home asset, and thus have no particular expertise over any foreign asset. We argue that to rationalize our novel empirical findings, a model needs to exhibit "partial" information specialization in equilibrium, so that individual investors find it optimal to produce information about some, but not all, foreign countries. We show that this can be achieved in an intuitive and parsimonious way by

[1] While we focus on sovereign debt because of the better data coverage over time, in Internet Appendix A.1 we show that sovereign debt exposures are highly correlated with both equity holdings and loans at the country level and are thus representative of the overall international portfolio allocation of European banks.

introducing some degree of "unlearnable" uncertainty to the standard set of assumptions of information models.

We begin our analysis by documenting two stylized facts about bank holdings of foreign sovereign debt, which we obtain from the European Banking Authority (EBA) stress tests. First, we find significant heterogeneity in the portfolio weights of individual foreign sovereign holdings, in the sense that banks significantly overweigh some foreign sovereigns, while underweighting others. This is a separate fact from the well-known home bias puzzle and speaks to concentration within the set of *foreign* portfolio holdings. Second, it is not the case that all banks put a high portfolio weight on the same sovereigns. Indeed, the correlation between the specific assets that individual banks overweigh is rather small. In other words, there is significant variation in the "affinity" that different foreign banks display for any given sovereign. Since the quarterly bond returns of different EEA sovereigns are only mildly correlated with one another (correlation of 0.54), at first blush this finding may suggest that banks are leaving potential diversification benefits on the table. Our key hypothesis, however, is that banks are not making such predictable mistakes in their investments, but rather take calculated risks in concentrating their portfolios in sovereigns for which they have precise and optimistic expectations of returns.

To test this hypothesis, we exploit the unique granularity of our data set and match the sovereign holdings of a bank to the same bank's forecasts of future sovereign yields. Our overarching finding is that portfolio shares of individual sovereign bonds are indeed sensitive to the characteristics of the underlying yield forecasts in ways that are both statistically and economically significant. Both higher expected quarterly returns and more precise forecasts predict larger portfolio holdings of a country's sovereign debt. In addition, as implied by optimal portfolio choice theory, there is a significant positive interaction between the two: the elasticity of portfolio shares to expected returns is increasing with the precision of the respective forecast. The results are highly significant, even after controlling for many other factors that could influence portfolios, such as time-varying bank- and country-specific unobserved factors, subsidiary structure, and cross-border mergers and acquisitions (M&A).

The positive interaction between the expected returns and the precision of the underlying forecasts is a particularly interesting and significant result. First, the magnitude of the interaction effect implies a risk aversion coefficient close to one, much lower than what is typically found for retail investors (Giglio et al. 2021). Second, it shows that the sophisticated investors in our sample indeed understand which of their forecasts are relatively more precise and take that into account in their portfolio allocations. Thus, our results are directly supportive of objective differences in information driving portfolio choice, as opposed to behavioral "familiarity" biases, which have been found to be important in the case of retail investors (e.g., Grinblatt and Keloharju 2001; Huberman 2001; Guiso et al. 2009).[2] Our interpretation of the results is not that banks possess special private information about the sovereigns that no other investor has, but rather that their macro research departments have developed different proprietary forecasting models, which results in heterogeneous expectations and ability to forecast future returns. In turn, the resultant comparative advantages in information play an important role in the international portfolio allocations of these institutional investors.

A potential concern is reverse causality, whereby banks could strategically (mis)report forecasts to rationalize their holdings or to move market prices in their favor. We address this concern in two ways. First, we emphasize that if banks systematically reported overly optimistic forecasts for the assets they hold, the resultant precision of those forecasts would be lower. As a result, the relation between precision and portfolio holdings would be the opposite of what we estimate. Second, we also consider a specification where we instrument bank forecasts with the average forecast of *nonbank* forecasters in Consensus Economics (mainly consulting firms and research institutes). These forecasts are unlikely to be driven by any bank-specific portfolio considerations or regulatory incentives and hence represent an exogenous forecast shifter for banks. The results remain very similar, and, in particular, our key estimate of the interaction term, which estimates how the effect of expected returns on portfolios changes with the precision of the forecast, remains virtually unchanged.

While our empirical findings are generally consistent with portfolio choice models with costly information production, our estimates put tight restrictions on the primitives of the information structure needed for such models to rationalize the data. Importantly, the information structure is a crucial degree of freedom for the theoretical literature, and its proper specification is still an open question. As Van Nieuwerburgh and Veldkamp (2010) show, different commonly used assumptions can make the model consistent with vastly different portfolio implications, ranging from full diversification to high concentration. With this in mind, Van Nieuwerburgh and Veldkamp (2009) exploit the presence of home bias in portfolios to argue in favor of a specification that leads to full specialization in information production about the home asset.

Our findings of significant heterogeneity in precision and related heterogeneity in portfolio holdings across individual *foreign* assets, however, are consistent with partial, but not full, information specialization, where investors also allocate costly information production resource to some, but not

[2] Our results do not imply that the forecasts themselves are optimal but that the forecasts are used appropriately in forming portfolios, given their statistical performance (which may not be perfect). In Internet Appendix A.2, we show that the survey forecasts in our sample indeed display many of the same apparent violations of rational expectations also found in other studies (Coibion and Gorodnichenko 2015; Bordalo et al. 2020).

all, foreign assets. This showcases how home bias, by itself, is not a sufficient statistic that can identify model ingredients.

We show theoretically that the needed partial specialization can be obtained parsimoniously by introducing some unlearnable uncertainty about asset payoffs. Specifically, we augment the baseline Van Nieuwerburgh and Veldkamp (2009) model with a two-factor asset payoff structure, where investors can produce additional information about only one of the factors. As a result, instead of displaying globally increasing returns, information production enjoys increasing returns initially, but diminishing returns past a saturation point. Thus, agents are incentivized to specialize in information production on a small set of assets, rather than on a singleton. We provide several analytical results that directly link the model's implications to our key empirical results, and show that the novel unlearnable uncertainty component is crucial for the model's ability to match the richer dimensions of the data that we document.

The paper contributes to the literature on international diversification, the bulk of which has focused on the classic home bias puzzle (Coeurdacier and Rey 2012). In this paper, we focus instead on the heterogeneity in portfolio allocations among individual foreign investments and the specific role played by information heterogeneity, utilizing a rich new micro-level data set with observations on both portfolio holdings and individual investor expectations.

Among other things, our results help differentiate between information models and behavioral familiarity based models (e.g., Grinblatt and Keloharju 2001; Huberman 2001; Guiso et al. 2009). Previously, these two classes of models have been difficult to distinguish between empirically. The complication arises from a lack of investor-asset-level panel data that reveal both portfolio holdings and expectations. The panel dimension is particularly important, as it allows for the computation of the objective precision of the underlying investor beliefs. This is crucial for differentiating between objective information advantages and behavioral familiarity, as we explain and showcase in the text.

For example, some previous studies have used indirect, and often aggregate, proxies of information frictions, such as cultural or geographic distance (e.g., Coval and Moskowitz 2001; Chan et al. 2005; Ivković and Weisbenner 2005; Guiso and Jappelli 2008; Leuz et al. 2009), but those can arguably be proxies for both objective differences in information and familiarity biases. Other studies have focused on aggregate capital flows (Ahearne et al. 2004; Portes and Rey 2005; Albuquerque et al. 2009; Curcuru et al. 2011) or the portfolio holdings of mutual funds (Gelos and Wei 2005; Hau and Rey 2008), and how they relate to destination country characteristics. Moreover, the literature that directly connects international portfolio shares and expected returns is small, and focuses on aggregate bilateral portfolio shares (Frankel and Engel 1984; Koijen and Yogo 2020) and flows (Bohn and Tesar 1996; Froot et al. 2001).

To the best of our knowledge, we are the first to analyze a data set that directly links the beliefs of individual institutional investors with their portfolio holdings at the individual asset level. For example, Stavrakeva and Tang (2020) link aggregate derivative positions to the average of survey of expectations of institutional traders, but do not observe beliefs and positions at the investor-asset level. Wang (2020) studies the link between the term structure of yield forecasts and U.S. banks' holdings of Treasuries, but does not analyze the accuracy of beliefs and its interaction with the point forecasts. On the other hand, many papers have focused on retail investors, where behavioral optimization mistakes and lack of sophistication play a key role (Campbell et al. 2007; Guiso and Jappelli 2008; Nechio 2014). For these types of investors, while the link between beliefs and portfolios is statistically significant, it appears to be quantitatively weak (Giglio et al. 2021). On the other hand, our findings indicate that sophisticated investors' portfolios are indeed closely aligned with beliefs heterogeneity.

On the theoretical side, we contribute to the literature that explores the interaction between information production and international portfolio choice (Nieuwerburgh and Veldkamp 2006; Van Nieuwerburgh and Veldkamp 2009; Mondria 2010; Dziuda and Mondria 2012; Valchev 2017). These models exploit the feedback between the information and portfolio choices to generate increasing returns in information, and thus full information specialization in home information production, which can generate significant information asymmetry between home and foreign assets as a whole. However, this full information specialization result is inconsistent with the heterogeneity across foreign portfolio holdings we document.

Our key innovation is to introduce a two-factor asset payoff structure where information can be produced about only one of the factors.[3] This weakens the feedback between portfolio and information choices, and leads to a model where investors optimally produce costly information for a subset of all assets, thus giving rise to the type of partial information specialization that rationalizes our empirical findings. We show that unlearnable uncertainty is necessary for the endogenous information framework to be able to match the richer dimension of the data we document, and, hence, it is a crucial component that should be incorporated in benchmark models going forward.

1. Data and Stylized Facts

We construct a data set of portfolio holdings and sovereign yield forecasts at the investor-asset-time level. To do so, we merge information on European banks' sovereign portfolios from the EBA stress tests with the forecasts the same banks report to the Consensus Economics survey. This yields a matched data set of holdings and forecasts at the bank level, offering a uniquely detailed view of institutional investors' portfolios.

[3] In doing so, we build on insights about information cost structures with eventually decreasing returns (as conjectured in Van Nieuwerburgh and Veldkamp 2010) and on the interaction between costly information acquisition and power utility (Peress 2004).

The EBA data, collected as part of bank stress tests and other regulatory exercises, come from a semiannual data set containing information on the asset holdings of large banks in the European Economic Area (EEA) from 2010Q4 to 2015Q2.[4] The number of banks included in each stress test varies between 60 and 82. The data set has particularly good coverage of bank sovereign debt holdings, as they are broken down by country of issuance and residual maturity. Other assets are reported less frequently. For example, the composition of equity and credit investments by country is available for 2012Q4, 2013Q2, and 2013Q4 only. In any case, foreign sovereign portfolio holdings are indicative of the overall geographic allocation of bank foreign assets, since sovereign debt holdings of a particular country are highly positively correlated with both equity and credit exposures to that country (see Internet Appendix A.1). Therefore, in our main set of results we focus on sovereign debt holdings.

One empirical challenge we face is that banks hold sovereign debt for a variety of reasons in addition to return maximization purposes. For example, sovereign debt has desirable liquidity and safety properties, since it can be used to meet regulatory requirements and can be pledged as collateral for private interbank or central bank refinancing operations. While in our regressions we control for a number of fixed effects that capture some of these factors, as a first step we also filter our data in several ways.

First, we restrict the sample to EEA sovereign debt holdings, because they receive the same regulatory treatment, that is, 0% risk weight (ESRB 2015), and are also all accepted at ECB refinancing facilities.[5] Moreover, we focus specifically on holdings of long-term bonds, defined as those with residual maturity of at least 5 years, as those are investment with significant future price risk. Importantly, those risks are not highly correlated, and, hence, these assets provide distinct investment opportunities. For example, the median (average) correlation of quarterly returns on long-term EEA sovereigns is only 0.56 (0.54). In addition, EEA sovereigns are the dominant form of security holdings for European banks, with a total of €2 trillion in EEA debt (7% of total assets), compared to €146 billion in equities.

Second, in the specific context of the European debt crisis, which is part of our sample period, additional factors may have contributed to bank holdings of domestic debt, such as moral suasion and financial repression (De Marco and Macchiavelli 2015; Ongena et al. 2019; Chari et al. 2020). Such factors could have distorted the choice of holding domestic debt and would thus affect the overall home bias in portfolios. Accordingly, we do not analyze the home bias itself, but rather focus on the foreign portion of sovereign portfolios and how the holdings of individual foreign sovereigns vary relative to one another.

[4] The stress tests were initially held at irregular intervals, and, thus, the following reporting dates are available: 2010Q4, 2011Q3, 2011Q4, 2012Q2, 2012Q4, 2013Q2, 2013Q4, 2014Q4, and 2015Q2.

[5] The criteria for the eligibility of assets as collateral for the Eurosystem credit operations are available online (https://www.ecb.europa.eu/mopo/assets/standards/marketable/html/index.en.html).

Third, and finally, we further restrict our attention to holdings that are marked-to-market (MTM) because they are kept in the available for sale, fair-value option, or held for trading accounting books, and remove from our baseline analysis any holdings that are classified as held-to-maturity (HTM). Only MTM assets, whose values in bank books must reflect day-to-day changes in market prices, can be sold off before maturity and are thus available for trading, while HTM assets typically are not. Overall, MTM assets are not bought to be passively held to maturity, and are thus more likely to be adjusted and traded as expectations of future prices and market conditions change. While roughly half of domestic sovereign holdings are held in the HTM book, the median share of foreign sovereign holdings held in the MTM book is 83%. The high percentage of foreign sovereign holdings accounted as MTM reassures us that these positions are indeed held for investment purposes.[6]

After constructing the portfolio data, we match the banks in the EBA sample to the forecasts these banks report to Consensus Economics, a survey of professional forecasters. The monthly forecasts are collected from a panel of forecasters at the beginning of each month and are published in the second week of the month. The panelists work for a variety of industry and research institutions, including banks' macro research departments. The forecasts are made by both domestic and foreign institutions for a set of key macroeconomic and financial variables of the major industrialized countries. In our analysis, we use forecasts of 10-year sovereign yields, which are the forecasts in Consensus Economics that are most relevant for computing the expected returns on long-term sovereign bond positions. We will explain the rationale underlying this approach next.

In particular, the survey provides us with 3-month-ahead forecasts of 10-year yields, namely, forecasts made in quarter t for the 10-year yield in quarter $t+1$. As bond yields and prices are inversely related, so are yield *forecasts* and *expected* quarterly bond returns. Abstracting from coupons (which are common knowledge and do not generate heterogeneity in expected returns across investors), the quarterly return on a 10-year bond is given by

$$r_{t+1} = 40 y_t^{(40)} - 39 y_{t+1}^{(39)} \approx 40 y_t^{(40)} - 39 y_{t+1}^{(40)},$$

where in the last expression we approximate the yield on a bond with maturity of 39 quarters with that of a 10-year bond. Our data give us bank-country-specific forecasts of $y_{t+1}^{(40)}$ and thus the quarterly expected return. To avoid the proliferation of superscripts, we write the forecast made in quarter t for the 10-year yield at $t+1$ as $F_{bt}(Y10_{ct+1})$, where b denotes the bank making the forecast and c is the respective country for which the forecast is made.

[6] The high MTM share also implies that our results are qualitatively unchanged if we use total foreign holdings, including the HTM portion, in our analysis. After an accounting rule change in May 2010 allowed banks to neutralize the effects of MTM unrealized gains or losses on bank regulatory capital and income, most (65%) sovereign bond holdings have been kept in the available for sale book.

Table 1
Summary statistics

Variable	Mean	SD	25th pct.	50th pct.	75th pct.	90th pct.	99th pct.	N
A. EBA foreign portfolio								
$Share_{bct}$	2.28	6.79	0.00	0.02	1.18	6.44	31.91	3080
$\mathbb{1}(Share_{bct})$	0.51	0.50	0.00	1.00	1.00	1.00	1.00	3080
$\mathbb{1}(ForeignFcst_{bct})$	0.13	0.34	0.00	0.00	0.00	1.00	1.00	3080
B. Consensus Economics								
$F_{bt}(Y10_{ct+1})$	3.37	1.48	2.2	3.5	4.35	5.06	7.6	8828
$SFE(Y10_{bct})$	0.35	0.58	.02	0.12	0.39	0.95	3.35	8815
C. EBA-Consensus Economics match								
$Share_{bct}$	6.72	12.56	0.59	2.76	8.78	16.59	71.12	212
$\overline{SFE}(Y10_{bc})$	0.48	0.28	0.29	0.40	0.60	1.03	1.43	212
$F_{bt}(Y10_{ct+1})$	3.37	1.67	2.13	2.91	4.50	5.78	7.80	212
log(dist)	6.48	0.60	6.03	6.32	7.04	7.22	7.27	212

This table provides summary statistics for all variables used in the empirical analyses. Panel A reports the summary statistics for the 16 EBA banks that report at least one foreign forecast on Consensus between 2006 and 2015. Panel B reports summary statistics for the full Consensus data set, including both banks and nonbanks, while panel C reports the summary statistics for the matched EBA-Consensus Economics data set. $Share_{b,c,t}$ is the share of foreign long-term (i.e., residual maturity > 5 years) MTM sovereign bond from EEA country c in bank b portfolio at time t. $\mathbb{1}(Share_{bct})$ is a dummy variable that equals one if $Share_{bct}$ is positive and zero otherwise. $\mathbb{1}(ForeignFcst_{bct})$ is a dummy variable that equals one if bank b makes a 10-year yield forecast about country c at time t and zero otherwise. $F_{bt}(Y10_{ct+1})$ is the 3-month-ahead forecast made by bank b regarding the 10-year yield on country c's sovereign debt in quarter t. $\overline{SFE}$(Y10) is bank b's average squared forecast error in forecasting the yield of country c. log(dist) is the natural logarithm of the population-weighted bilateral distance measure of (Mayer and Zignago 2011) between foreign country c and the country of bank b's headquarter.

Furthermore, we use the fact that we have a time series of individual forecasts to construct an objective measure of the ex post precision of the forecasts at the bank-country level. We compute the precision of bank b's forecast for country c as the average squared forecast error (SFE) for the full period available from Consensus (2006–2015):

$$\overline{SFE}_{bc} = \frac{1}{T}\sum_{t=1}^{T}(F_{bt}(Y10_{ct+1}) - Y10_{c,t+1})^2,$$

where $Y10_{c,t+1}$ is the actual realization of the 10-year yield of country c in quarter $t+1$.

The yield forecasts from Consensus Economics are available monthly from September 2006 to December 2015 for 11 different EEA countries and for a worldwide sample of 85 banks. Of them, 36 are EBA banks for which we have detailed portfolio holdings data and 16 of them have made at least one forecast for a *foreign* EEA country. The portfolios and forecasts of those 16 banks, mostly large global institutions, form our baseline estimation sample. Overall, those banks account for 50% of all EEA sovereign debt holdings in the EBA database—over €1 trillion—providing us with good coverage of aggregate debt holdings.

Table 1 displays summary statistics on the portfolios of EBA banks who report foreign forecasts in Consensus (panel A), all yield forecasts available in Consensus by both banks and nonbanks (panel B) and the matched bank-country-quarter data set for which we have both portfolio data and a yield forecast (panel C). We discuss these summary statistics next.

1.1 Portfolio characteristics

While portfolio home bias, the primary focus of the previous literature, is also a feature of our data, in this section we bring attention to two different stylized facts about foreign portfolio holdings.[7] First, there is significant heterogeneity across the portfolio weights assigned to individual foreign assets, in the sense that banks significantly underweight some foreign assets, while overweighting others relative to their respective market weights (similar to the evidence on equity holding of mutual funds in Chan et al. 2005). As a result, we essentially find that portfolios are concentrated not only because of the classic home bias phenomenon but also because the foreign portion of portfolios is concentrated as well. Second, not all banks tend to under- and overweigh the same foreign sovereigns, but rather there seems to be bank-specific variation in which countries are over- and underweighted.

To facilitate a comparison with the previous literature, we measure over- and underweighting relative to the standard capital asset pricing model (CAPM) benchmark (i.e., market portfolio weights). This is not to say that the CAPM is the optimal portfolio, and, hence, the deviations we find are "suboptimal." In fact, our own model and empirical analysis imply that the deviations are due to the optimal exploitation of comparative advantages in information production. Instead, our goal in this first section is to simply show that banks' portfolio holdings of individual foreign assets are highly heterogeneous in way that is relatable to the previous literature, which similarly looks at home portfolio shares standardized by market size. In our main regression analysis in Section 2 we do not standardize portfolios in any way, but rather control for a variety of potential factors via a rich set of fixed effects and other variables.

To illustrate the extent of portfolio heterogeneity, we construct a "foreign-bias" index for each bank-destination-country pair in each quarter (suppressing time subscripts),

$$FB_c^b = \tilde{x}_c^b - \tilde{x}_c^{mkt}$$

where $\tilde{x}_c^b = x_c^b / \sum_{c \neq H} x_c^b$ is the portfolio weight of foreign asset c ($c \neq H$, where H is the index for the country of domicile of bank b) as a share of the total foreign investment of bank b, and $\tilde{x}_c^{mkt} = x_c^{mkt} / \sum_{c \neq H} x_c^{mkt}$ is the market value of country c sovereign debt as a share of the market value of all EEA sovereign bonds except for EEA debt that is domestic to bank b. In this notation, x_c^b is the level of the holdings of country c sovereign bonds by bank b. For subsidiaries, the "home" country is that of the headquarters of the parent bank.

[7] For details on the portfolio home bias in our data set, please see Internet Appendix A.4.

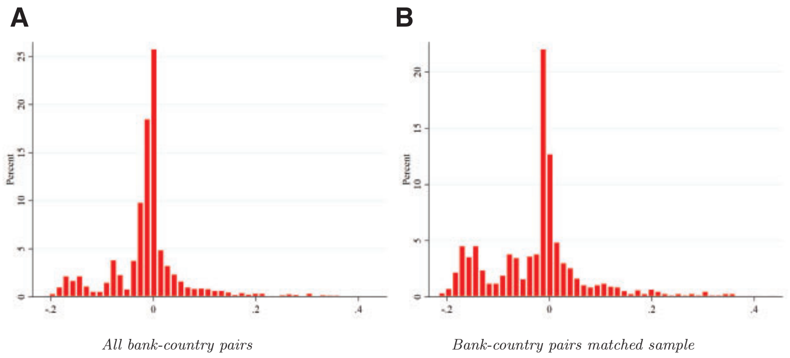

Figure 1
Empirical foreign bias distribution
Panel A plots the distribution of the foreign bias index across all bank-EEA country pairs available in the EBA data set (23 countries) for the 16 banks in the matched EBA-Consensus data set. Panel B plots the distribution of the foreign bias index only for the 11 countries for which we also have Consensus forecasts for the same set of 16 banks. The distribution has been winsorized at the 1st through 99th percentiles.

This foreign bias index varies at the bank-country level and equals zero when a bank's investment in country c equals the respective market weight of country c's debt as a share of the total outstanding EEA debt. A positive value corresponds to overweighting of that asset and a negative value to underweighting it, compared to the foreign market portfolio. Overall, this statistic can help us understand to what extent portfolio holdings of individual foreign assets are more or less concentrated *relative to each other*. We stress that values different from zero do not mean that the holdings are necessarily suboptimal. Rather, using the market portfolio as the benchmark provides us with a useful statistic that is comparable across assets in both low and high supply, and connects us with the previous literature (Chan et al. 2005; Coeurdacier and Rey 2012).

Panel A of Figure 1 plots the empirical distribution of the foreign bias index. This distribution is centered on zero, with a median value of −0.006, which means that individual asset holdings tend to be underweighted roughly as often as being overweighted. Importantly, the distribution features a large amount of dispersion: the standard deviation of FB_c^b is 0.087, meaning that an "overweighted" asset with FB_c^b one standard deviation above the median has a portfolio share that is 8.7 percentage points larger than its corresponding market share. This is a very large number, given that the median market share of an EEA sovereign is 2%. Moreover, it is not the case that all banks tend to under- and overweight the same assets. The correlation of FB_c^b across banks and countries is only 0.23.

We also find that the heterogeneity in FB_c^b is even larger once we condition on the subset of EEA countries for which we observe banks making forecasts in the Consensus survey (11 countries in total). Panel B of Figure 1 plots this distribution. It is similarly centered on a value close to zero (the median is −0.012), but displays a standard deviation of 0.106, 22% larger than in the overall sample of panel A. This implies that portfolio heterogeneity is especially pronounced in the case of countries for which banks feel confident enough to report yield forecasts to Consensus.

Given the relatively low correlation in the quarterly bond returns of EEA sovereigns in our sample—median (average) correlation of just 0.56 (0.54)—one interpretation of this evidence is that banks are leaving a lot of potential diversification benefits on the table.[8] Our hypothesis, however, is that banks are not making simple investing mistakes, but rather take calculated risks with their investment portfolio, and concentrate positions in sovereigns for which they have not only precise but also particularly optimistic expected returns. To motivate this hypothesis further, in the next subsection we showcase that the asset-specific forecasts are indeed highly heterogeneous as well.

1.2 Forecast characteristics

Table 1, panel B, shows that the Consensus Economics forecasts are quite accurate, with a mean squared forecast error (SFE) of just 35 basis points (10 times lower than the unconditional variance of yields of 3.69%). However, there is significant dispersion in forecast precision: the standard deviation of SFE is 0.58%, yielding a coefficient of variation of 1.66. A similar pattern emerges in the matched EBA-Consensus sample (panel C), suggesting that the banks in our sample have a significant, but heterogeneous ability to forecast future yields.

Panel A of Figure 2, plots the histogram of yield forecasts across banks and countries. To visualize them all in one histogram, we standardize forecasts by the average yield forecast for each country in each month, to control for differences in the level of yields across countries. The resultant distribution gives us a sense of the baseline heterogeneity in forecasts across banks. And indeed the dispersion is significant. The distribution of point forecasts has a standard deviation of 0.63% and an interquartile range of 0.68%.

There is also significant heterogeneity in the precision of the forecasts, both across banks and across the different country forecasts a single bank makes. Panel B of Figure 2 plots the histogram of $\overline{SFE}_{bc}$ standardized by the average $\overline{SFE}_{bc}$ at the country level. Despite a high degree of accuracy in the forecasts on average, there are substantial differences in the ability of different banks to forecast any given yield. The standard deviation of the above distribution is 0.22% with an interquartile range of 0.21%.[9]

---

[8] We have also evaluated the implied diversification benefit more directly, by comparing the observed bank portfolios with the minimum-variance portfolio implied by the empirical variance-covariance matrix of bond returns. We find that the average (median) bank portfolio exhibits 128% (99%) higher variance.

[9] Banks also display home bias in information: home forecasts are more precise than foreign forecasts. Since we are primarily interested in the differences across individual foreign investments, we relegate the information home bias analysis to Internet Appendix A.5.

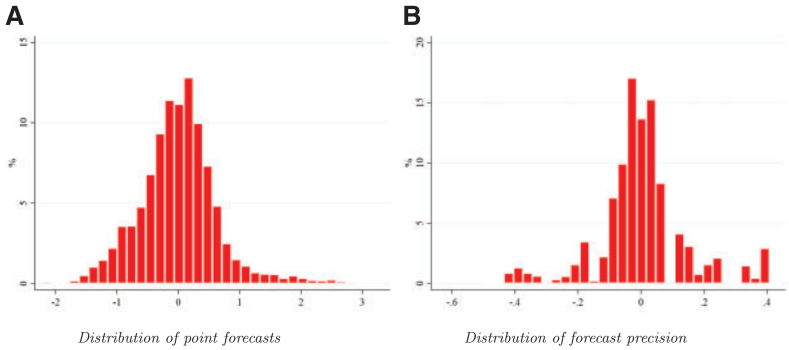

Figure 2
Information heterogeneity
Panel A plots the distribution of 10-year yield forecasts minus the country-month average (i.e., the residual from a regression of 10-year yield forecasts $\widehat{Y10}_{b,c,t}$ on country-month fixed effects). Panel B plots the distribution of bank-country-specific squared forecast errors $\overline{SFE}_{bc}$ minus the country average (i.e., the residual from a regression of $\overline{SFE}_{bc}$ on country fixed effects).

The reader may wonder whether the forecasts communicated to Consensus by the macro research teams are (a) truly representative of the bank's actual expectations and (b) taken into account by the trading desk of the bank. Regarding the first concern, banks have an incentive to report their best forecasts to Consensus to build a reputation as accurate forecasters. For example, Consensus gives an annual "Forecast Accuracy Award," publishing the names of the research economists who produced the most accurate forecasts for each country, along with the name of the institution they work for. This award is valued by the banks, which often release press statements upon receiving it.[10] As to the second concern, macro research teams regularly communicate their views to the bank's rate strategists and traders. The so-called "Chinese walls," which typically exist between traders and the M&A department, are not present between the macro research teams and the bond trading desk, as evidenced, for example, by standard job posts for macro research positions, which highlight that the economist will support investment decisions.[11] Moreover, there is a time lag of about 2 weeks between the time the survey is taken and when it is published, which gives traders enough time to implement their trades before the information becomes public, and thus provides further incentive for the bank to submit their best forecasts.

More directly, we show that the forecasts released to Consensus indeed contain valuable information in predicting future yields. To illustrate this point, first we estimate an AR(1) forecasting model for yields augmented with the

10 See, for example, Intesa San Paolo and Societe Generale.

11 For example, a February 2021 online job posting by Goldman, Global Investment Research states "from macroeconomic forecasts to individual stock analysis, our team develops tools and insights to help shape investment strategies for clients and the firm," the firm being the bank itself.

Table 2
Predictability of sovereign yields

	(1)	(2)	(3)	(4)	(5)	(6)
$Y10_{c,t}$	0.98***	0.95***	0.83***	0.16***	0.13**	0.05
	(0.01)	(0.02)	(0.03)	(0.05)	(0.05)	(0.06)
$\overline{F_t(Y10_{c,t+3})}$				0.90***	0.93***	0.95***
				(0.06)	(0.06)	(0.07)
$H_0: \gamma_1 = 1$ (p-val)	(.14)	(.01)	(.00)			
$H_0: \gamma_2 = 1$ (p-val)				(.09)	(.20)	(.46)
Country FE	no	yes	yes	no	yes	yes
Time FE	no	no	yes	no	no	yes
N	1,107	1,107	1,107	1,107	1,107	1,107
R^2	.92	.92	.96	.95	.95	.98

This table provides the estimates for Equation (1). The dependent variable is $Y10_{c,t+3}$, the realized 3-month-ahead (one quarter) 10-year sovereign yield of country c, $Y10_{c,t}$ is the realized 10-year yield for country c in month t and $\overline{F_t(Y10_{c,t+3})}$ is the average yield 3-month-ahead forecast made in month t across all forecasters that report 10-year yield forecasts for country c in Consensus Economics. Robust standard errors are reported in parentheses. *p <.1; **p <.05; ***p <.01.

average Consensus forecast:

$$Y10_{c,t+3} = \gamma_1 Y10_{c,t} + \gamma_2 \overline{F_t(Y10_{c,t+3})} + \alpha_c + \alpha_t + \varepsilon_{c,t+1} \tag{1}$$

where $Y10_{c,t+3}$ is the realized 3-month-ahead 10-year sovereign yield of country c, $\overline{F_t(Y10_{c,t+3})}$ is the average 3-month-ahead yield forecast for country c made in month t across forecasters in Consensus Economics, $Y10_{c,t}$ is the realized 10-year yield for country c in month t, α_c is a country-specific intercept, and α_t is a month fixed effect. Table 2 reports the estimates.

As expected, sovereign yields are highly autocorrelated. In the stand-alone AR(1) specification, the coefficient γ_1 is 0.98 and not significantly different from one in the specification with no fixed effects, and remains high (0.83) even when we include country and month fixed effects. However, when we control for the average Consensus forecast in columns 4 to 6, the AR(1) coefficient becomes much smaller and statistically insignificant with country and time fixed effects. On the other hand, our estimate for γ_2 is above 0.9, significantly different from zero at the 1% level and not significantly different from one. Thus, the Consensus forecasts contain significant information above and beyond that contained in past asset prices.

In fact, the banks' Consensus forecasts can help generate excess returns. To show this, we compute the profits of the following trading strategy. Each month we sort sovereign bonds based on the expected quarterly return as implied by the difference between the current yield and the average yield forecast of our EBA banks, and then go long in the top third and short in the bottom third of this sorted list. This strategy yields an (annualized) excess return of 5.21% with a Sharpe Ratio of 0.72. As a comparison, a classic high-vs-low yield strategy that goes long in high-yield "peripheral" countries, and short in low-yield "core" countries, yields an excess return of just 0.98% with a Sharpe ratio of 0.09. Thus, the forecasts submitted to Consensus contain valuable information that can generate excess returns. In the same spirit, our basic hypothesis is that banks use their own forecasts when forming portfolios.

Lastly, we stress that our hypothesis is not that individual banks have exclusive or private knowledge about any given sovereign that no other investor has. On the contrary, we argue that they have differential abilities to extract insights from public data for the purpose of making forecasts, leading to heterogeneous beliefs about future foreign yields. This is likely due to the different proprietary forecasting models that are used to produce the forecasts within each bank, as well as differences in the overall choice of allocating scarce analytical resources in forecasting any specific country. In terms of terminology, we refer to the heterogeneity in forecast characteristics as "information heterogeneity," since the effective information that the forecasting teams of the different banks extract from the public data is different.

2. Forecast Heterogeneity and Portfolio Holdings

The stylized facts presented in the previous section highlight the heterogeneity in both foreign portfolios and forecasts. While in Section 3 we provide a general equilibrium model where information, beliefs, and portfolio heterogeneity are all endogenous, it is useful to first describe the empirical link between portfolio and forecast heterogeneity with a reduced-form analysis, whose intuition can be derived from a simple partial equilibrium framework. For example, in the Merton (1969) model with a diagonal variance-covariance matrix of returns, the optimal portfolio share of a risky asset c is given by

$$Share_{c,t} = \frac{\mathbb{E}_t\left(R_{c,t+1}\right) - R_f}{\gamma \operatorname{Var}(R_{c,t+1})}, \tag{2}$$

where $\mathbb{E}_t(R_{c,t+1})$ is the investor's expected return on risky asset c, $\operatorname{Var}(R_{c,t+1})$ is the investor's perceived variance of that return, γ is the coefficient of relative risk aversion, and R_f the risk-free rate. The key forces are intuitive. Portfolio holdings are increasing in expected returns and decreasing in the variance of returns. Importantly, there is also an interaction effect: portfolio holdings are more sensitive to expected returns when the perceived variance of returns is lower. In that case, investors bet in the direction of their expected returns more aggressively.

To test these relationships in the data, we estimate the following regression:

$$\begin{aligned} Share_{bct} = {} & \beta_1 \overline{SFE}_{bc} + \beta_2 F_{bt}(Y10_{c,t+1}) \\ & + \beta_3 \overline{SFE}_{bc} \times F_{bt}(Y10_{c,t+1}) + \lambda \ln(Dist)_{bc} + \mu_b + \gamma_{ct} + \varepsilon_{bct}, \end{aligned} \tag{3}$$

where $Share_{bct}$ is the share of sovereign bonds with residual maturity above five years issued by country c in the portfolio of bank b, where b is headquartered in a country other than c.[12] $F_{bt}(Y10_{c,t+1})$ is bank b's one-quarter-ahead forecast

12 We focus on long-term bond holdings since our forecast data are for 10-year yields; however, the results are robust to using all sovereign holdings, including short-term debt (Internet Appendix Table A6).

of the 10-year yield of country c made in quarter t.[13] Because of the inverse relationship between yields and bond prices, the yield forecast is a direct, but inverse proxy for the expected return, $\mathbb{E}_t(R_{c,t+1})$. In turn, the bank-country-specific squared forecast error of the yield forecasts, $\overline{SFE}_{b,c}$, provides us with a direct estimate of the investor's perceived uncertainty in her forecast of returns, that is, $\text{Var}(R_{c,t+1})$. This measure of forecast uncertainty is different from the measure of forecast dispersion typically found in the literature studying the predictability of returns in the US Treasury market (Giacoletti et al. 2021). We do not measure disagreement among forecasters, but rather each forecaster's individual ability to forecast a specific sovereign yield.

We also control for the well-known "gravity" relation in international capital flows (e.g. Portes and Rey 2005) by including the log of the bilateral distance, $\ln(Dist_{bc})$, between the country of the bank's headquarter and the country of investment, weighted by population size (Mayer and Zignago 2011). Finally, μ_b and γ_{ct} are bank and destination country-time fixed effects, which among other things, control for the time-varying credit risk, liquidity, and yield of each sovereign. We cluster standard errors at the bank-destination country level given that the key variation we exploit for identification is a bank's forecast and its precision for specific foreign countries.[14]

Table 3 presents the estimates. First of all, we find that distance, when included by itself, is indeed negatively related to portfolio holdings. In the most saturated gravity specification with bank and country-time fixed effects (column 3), the estimated coefficient implies that doubling the distance between the home and foreign country is associated with a 6.83 percentage points decrease in the portfolio share of that foreign country's sovereign. That is a sizable effect considering that the mean share of long-term foreign sovereign holdings in the matched sample is 6.72%. However, the effect of geographical distance disappears once we include our forecast-based information variables (columns 4 to 6), suggesting that the presence of gravity in international capital flows is indeed related to information differences.

This is a novel result: because of data limitations, existing studies only have been able to relate distance to indirect proxies for information such as cross-listing, corporate governance, and disclosure rules (Ahearne et al. 2004;

[13] While forecasts are monthly, the portfolio holdings data are reported semiannually, either on June 30 or on December 31 (the only exception being September 30, 2011). To merge the two data sets, we take the average of the forecasts made in the quarter of the holdings data. For example, if the holdings data refer to December, we take the average of the 3-month-ahead forecasts in October, November, and December. We do so to maximize data coverage, since not all banks report forecasts in December, but they may do so in November or October. Finally, we note that all banks are surveyed in the beginning of the month, and, hence, even the December forecast is made before the date the portfolio data refers to.

[14] It is worth mentioning that the estimating sample for Equation (3) is only a fraction of our entire data set, since we lose a substantial amount of observations in matching portfolios and forecasts. While most banks report a forecast for the home country's yield, only the large global banks (16 in our sample) forecast multiple foreign countries' yields, with a median number of foreign forecasts of 3 and a maximum of 6 (40 bank-country clusters). Given the small sample size, one alternative to clustering is to use bootstrapped standard errors. We do so in Internet Appendix Table A7 and show that statistical significance is similar.

Table 3
Information heterogeneity and portfolio holdings

	(1)	(2)	(3)	(4)	(5)	(6)
log(dist)	−2.87	−3.43*	−6.83**	−0.18	−1.70	−3.31
	(2.45)	(1.80)	(3.28)	(2.02)	(1.83)	(3.39)
$\overline{SFE}$(Y10)				−36.11**	−33.40***	−35.96***
				(13.87)	(10.61)	(13.17)
$F(Y10)$				−4.31**	−3.18*	−5.50*
				(2.01)	(1.62)	(2.80)
$\overline{SFE}(Y10) \times F(Y10)$				7.59**	6.57**	9.05***
				(3.58)	(2.87)	(3.17)
Bank FE	No	Yes	Yes	No	Yes	Yes
Time FE	No	Yes	–	No	Yes	–
Country-time FE	No	No	Yes	No	No	Yes
Observations	212	211	188	212	211	188
R^2	.02	.45	.74	.12	.52	.77

This table provides the estimates for Equation (3). The dependent variable is the share of foreign EEA country c sovereign bonds in bank b sovereign portfolio of long-term debt (> 5 years residual maturity). log(dist) is the natural logarithm of the population-weighted bilateral distance measure of (Mayer and Zignago 2011). $\overline{SFE}$ is bank b's average squared forecast error on 10-year yields of country c. $F(Y10)$ is the 3-month-ahead forecast made by bank b regarding the 10-year yield on country c's sovereign. Standard errors are reported in parentheses and are clustered at the bank-country level. $^{*}p$ <.1; $^{**}p$ <.05; $^{***}p$ <.01.

Leuz et al. 2009) or telephone calls and the number of multinational bank branches (Portes and Rey 2005). Such indirect proxies, however, cannot differentiate the information explanation from behavioral familiarity (Grinblatt and Keloharju 2001; Huberman 2001). Our data allow us to confirm the information hypothesis directly.[15]

Importantly, the estimated coefficients for the forecast-based variables ($\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$) are statistically significant, even when we absorb country shocks in each time period with country-time fixed effects, and align with the predictions of standard portfolio theory, as illustrated in Equation (2). Specifically, portfolio holdings are increasing in expected returns, that is, lower future yields ($\beta_2 < 0$), and decreasing in perceived uncertainty ($\beta_1 < 0$).[16] We also find that the elasticity of portfolio holdings to expected returns is increasing in the precision of the specific bank-country forecast ($\beta_3 > 0$).

The estimated coefficients are also economically significant. The estimates in column 6 of Table 3 indicate that a one-standard-deviation decrease in the expected future yield (1.67%) for a forecast with median precision ($\overline{SFE} =$ 0.40%) is associated with a portfolio share increase of 3.14 percentage points, which is roughly half of the average portfolio share of foreign sovereigns.

15 In unreported results (available on request), we find that information variables are indeed directly correlated with distance. Banks are less likely to make a forecast about countries farther away from their headquarter and make less precise forecasts about them if they do make a forecast.

16 Drawing a distinction between our measure of investor-level uncertainty, $\overline{SFE}_{bc}$, and disagreement across forecasters at a point in time is helpful. While recent work has shown that disagreement, or dispersion of forecasts, is useful in forecasting Treasury excess returns (Giacoletti et al. 2021), that measure is not related to the heterogeneity in portfolio holdings in our data (see Internet Appendix Table A8). This further confirms that our results are due to the link between banks' portfolios and their own forecasts, not to overall disagreement among forecasters.

However, the same one-standard-deviation increase in expected returns based on a more precise forecast (25th percentile $\overline{SFE}=0.29\%$) is associated with a portfolio share increase of 4.8 percentage points, an effect that is more than 50% stronger.

Overall, the estimates imply an economically significant elasticity of portfolio holdings to expected returns. To conceptualize these numbers, we perform a back of the envelope calculation using Equation (2) and compute the coefficient of risk aversion γ implied by our estimates. Given the high precision of beliefs (median $\overline{SFE}=0.40\%$) the estimates from column 6 imply $\gamma=1.33$, which is a reasonable and in fact relatively low level of risk aversion. For example, it contrasts with the high risk aversion ($\gamma=36$) implied by data on retail investors (Giglio et al. 2021). Thus, the high elasticity of holdings to expected returns we find is a reflection of the high precision of the banks' forecasts of future bond returns.

Moreover, the large and positive $\widehat{\beta}_3$ estimate suggests that banks internalize which of their country-specific forecasts are most precise, and put a higher weight on the more precise forecasts when forming portfolios. This implies that the concentration that we find within foreign portfolios is indeed a seemingly rational response to objective differences in forecast quality, rather than a phenomenon driven by behavioral motives, such as familiarity. Specifically, while a behavioral agent might also have more precise information about familiar assets ($\beta_1<0$), she is unlikely to display a higher portfolio sensitivity to changes in point forecasts that are objectively more precise ($\beta_3=0$ for a behavioral investor).[17]

2.1 Robustness tests

2.1.1 Richer fixed effects structure. We acknowledge that while our benchmark specification controls for time-varying country factors and factors that are fixed at the bank level, there could be considerations that are bank specific but vary over time (e.g., bank-specific shocks to the balance sheet) and even factors that are fixed at the bank-destination-country level (e.g., international subsidiaries structures). Our data set is rich enough to allow us to control for all of these factors via bank-time and bank-country fixed effects. We reestimate our main regression in Equation (3) with this richer set of fixed effects and present the results in columns 1 to 3 of Table 4. These specifications involve some loss of observations, but the key coefficients of interest, particularly the interaction term β_3, remain statistically significant and

[17] While banks put more weight on objectively precise forecasts, this portfolio rationality does not necessarily imply that the forecasts themselves are fully optimal. As shown in previous research, survey forecasts often violate the paradigm of rational expectations, and our survey is no exception. We find evidence of extrapolation (Gennaioli et al. 2016) and overreactions to macroeconomic news at the individual forecaster level (Bordalo et al. 2020; Wang 2020). See Internet Appendix A.2 for details. Our point is that, given the statistical properties of forecasts, banks appear to internalize and exploit the emerging comparative advantages in information when forming portfolios.

Table 4
Information heterogeneity and portfolio holdings: Robustness

	Bank-time, bank-country FE			Cross-border M&A		
	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{SFE}$(Y10)	−31.84*			−36.45**	−45.73**	
	(16.43)			(13.82)	(21.12)	
$F(Y10)$	−1.03	−5.98*	−6.17*	−5.48*	−6.11	−6.51**
	(3.47)	(3.35)	(3.37)	(2.80)	(4.83)	(2.90)
$\overline{SFE}(Y10) \times F(Y10)$	9.32*	8.18**	11.06**	9.10***	11.88**	11.59**
	(4.69)	(3.77)	(4.97)	(3.23)	(5.05)	(5.53)
$CrossBorder\ M\&A_{bct}$				0.69		2.72
				(2.31)		(20.21)
Country-time FE	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	–	–	–	Yes	Yes	-
Bank-time FE	Yes	No	Yes	No	No	Yes
Bank-country FE	No	Yes	Yes	No	No	Yes
Observations	132	181	125	188	139	125
R^2	.86	.83	.92	.77	.78	.92

Robustness tests for the estimates in Table 3. Columns 1–3 include bank×time, bank×destination country or both fixed effects together. Columns 4 and 6 control for $CrossBorder\ M\&A_{bct}$, a dummy variable equal to one for all periods after bank b engages in a cross-border M&A with a bank located in foreign country c in year t. In column 5 we exclude from the sample all banks that engage in a cross-border M&A at any point in time during the sample period. Standard errors are reported in parentheses and are clustered at the bank-country level. *p <.1; **p <.05; ***p <.01.

quantitatively very similar to the benchmark estimates in Table 3. Thus, our results are robust to a variety of additional unobserved heterogeneity.

2.1.2 M&A activity. The additional bank-time and bank-country fixed effects would still not necessarily capture mergers and acquisitions (M&A) activity that happens during our sample period. Consolidations could indeed introduce simultaneous shifts in portfolio shares and information acquisition. We think that this is unlikely, due to our focus on foreign portfolios and the fact that the bulk of consolidations in our sample period happened domestically. Nevertheless, to fully account for the effect of such mergers and acquisitions, we have collected information about cross-border M&A for the banks in our sample. Utilizing this new data, we estimate an augmented version of our main regression that includes an indicator variable ($CrossBorder\ M\&A_{bct}$) that takes the value of one if bank b has acquired a bank in country c for all time periods t following the merger. The banks in our regression sample have only 12 M&A events, 5 of which involve cross-border acquisitions by a total of 3 banks in our sample. Column 4 in Table 4 shows that including $CrossBorder\ M\&A_{bct}$ as a control does not affect our results with our benchmark set of fixed effects. Column 6 shows that this remains the case even with the most stringent possible set of fixed effects. Moreover, in column 5 we exclude the three banks that engaged in any cross-border M&A from the sample and the results again change very little.

2.1.3 IV approach. A potential concern with the results of Table 3 is reverse causality between portfolio shares and banks' sovereign yield forecasts. This could happen for mainly two reasons. First, a bank may report overly optimistic forecasts in order to induce other investors, who observe the forecasts when they become public on Consensus (after a 2-week lag), to purchase the asset, while the bank itself is selling it. We think that this is unlikely to be the case because then higher expected returns (i.e., lower future yield forecasts) would be associated with smaller holdings ($\beta_2 > 0$), which is the opposite of what we find. Alternatively, a bank may systematically inflate predictions and report overly optimistic forecasts for the assets it holds in large quantities, to encourage others to buy, leading to higher capital gains for the bank. This is also unlikely because systematically producing overly optimistic forecasts for large holdings would create higher squared forecast errors for those sovereigns, and thus generate $\beta_1 > 0$ and $\beta_3 < 0$, the opposite of what we find.

Nevertheless, to address the potential reverse causality concern more directly, we consider an instrumental variable (IV) approach for a bank b's forecast using the average forecast for country c's future yield made in quarter t by all the nonbank forecasters in Consensus Economics domiciled in countries other than that of bank b. These nonbank forecasters are mostly economic consulting firms and research institutes and, as such, are not subject to the same incentives that may influence bank portfolio decisions and hence the forecasts of European banks. If banks in our sample are reporting forecasts in part to rationalize their holdings, the instrument serves as a forecast shifter that is uncorrelated to the foreign portfolio holdings of any individual bank, but is correlated with bank forecasts to the extent that there is a common component in forecasts.[18]

Formally, the model of Equation (3) has two endogenous variables: $F_{bt}(Y10_{ct+1})$ and $\overline{SFE}_{bc} \times F_{bt}(Y10_{ct+1})$.[19] We thus construct two instruments, $F_{bt}^{nonbank}(Y10_{ct+1})$ and $\overline{SFE}_{bc} \times F_{bt}^{nonbank}(Y10_{ct+1})$, and run two first stage regressions as follows:

$$F_{bt}(Y10_{ct+1}) = \theta_1 F_{bt}^{nonbank}(Y10_{ct+1}) + \theta_2 \overline{SFE}_{bc} \times F_{bt}^{nonbank}(Y10_{ct+1}) \qquad (4)$$
$$+\alpha_1 \ln(Dist)_{bc} + \alpha_2 \overline{SFE}_{bc} + \mu_b + \gamma_{ct} + \varepsilon_{bct}$$

[18] To construct the IV, we require at least three nonbank forecasters for each observation. Only a few small countries (Netherlands, Norway, Slovakia, and Sweden) have less than three nonbank forecasters in some quarters. In these cases, in our IV we also include forecasts from banks that are domiciled in countries other than country c and bank b's home country. We do so under the assumption that the sovereign portfolios of banks located in other countries are less likely to be correlated with bank b's specific exposure to a foreign country. In line with the same logic, we also exclude the forecasts made by local foreign banks, that is, banks domiciled in the foreign country c, which, given the home bias in portfolios, may also produce forecasts to be in line with their holdings.

[19] We do not instrument the stand-alone regressor $\overline{SFE}_{bc}$ even though it contains the endogenous point forecasts. This is because *any* endogenous bias in $F_{bt}(Y10_{ct+1})$, either positive or negative, would have the same effect of inflating $\overline{SFE}_{bc}$, resulting in no systematic correlation between portfolios and any potentially biased $\overline{SFE}_{bc}$. This works like classical measurement error, which leads to an attenuation of the SFE-associated coefficients, β_1 and β_3. Therefore, our estimation results of $\beta_1 < 0$ and $\beta_3 > 0$, remain conservative.

Table 5
Information heterogeneity and portfolio holdings: Foreign nonbank IV

	OLS	1st stage		2nd stage
		$F(Y10)$	$\overline{SFE}(Y10)\times F(Y10)$	
	(1)	(2)	(3)	(4)
$\overline{SFE}$(Y10)	−35.96***	0.21	−0.03	−41.79**
	(13.17)	(0.38)	(0.30)	(18.24)
$F(Y10)$	−5.50*			−17.69***
	(2.80)			(5.83)
$\overline{SFE}(Y10)\times$ F(Y10)	9.05***			10.48**
	(3.17)			(4.19)
$F(Y10)_{nonbank}$		−3.68***	−3.28***	
		(0.50)	(0.33)	
$\overline{SFE}(Y10)\times F(Y10)_{nonbank}$		−0.08	0.96***	
		(0.08)	(0.07)	
Bank FE	Yes	Yes	Yes	Yes
Country-time FE	Yes	Yes	Yes	Yes
Observations	188	188	188	188
1st-stage F-stat				26.19

This table provides first- and second-stage estimates for Equation (3). The dependent variable in the second stage is the share of EEA country c sovereign bonds in bank b sovereign portfolio of long-term debt (>5 years residual maturity). $F(Y10)$ is the 3-month-ahead forecast made by bank b regarding the 10-year yield on country c's sovereign debt and it is instrumented by $F(Y10)_{nonbank}$, the average forecast on country c's 10-year sovereign yield made by all nonbank forecasters in that quarter domiciled in other countries than bank b. The 1st-stage F-stat is the Kleibergen-Paap F-statistic for weak identification. Standard errors are reported in parentheses and are clustered at the bank-country level. *p <.1; **p <.05; ***p <.01.

$$\overline{SFE}_{bc}\times F_{bt}(Y10_{ct+1})=\theta_3 F_{bt}^{nonbank}(Y10_{ct+1})+\theta_4\overline{SFE}_{bc}\times F_{bt}^{nonbank}(Y10_{ct+1})+\alpha_3\ln(Dist)_{bc}+\alpha_4\overline{SFE}_{bc}+\mu_b+\gamma_{ct}+\varepsilon_{bct} \tag{5}$$

The results of the first- and second-stage estimates are presented in Table 5. The two first-stage equations (columns 2 and 3) show that the IV is strongly correlated with the endogenous variable, since the estimates for θ_1,θ_2, and θ_4 are highly significant and overall yield a Kleibergen-Paap first-stage F-stat of 26, above conventional levels.[20]

Moreover, the main results remain the same, and perhaps are even stronger, as the stand-alone effect of yield forecasts on portfolios increases. Importantly, the interaction coefficient β_3 in column 4 remains positive, significant, and similar in magnitude to the ordinary least squares (OLS) estimate in column 1. Thus, the IV results suggest that, even if the OLS estimate of β_2 may be attenuated, the crucial interaction coefficient β_3 seems not to be affected by reverse causality.

[20] We note that the negative first-stage coefficient on $F(Y10)_{nonbank}$ is simply a product of our rich set of fixed effects. Specifically, when we include country-time fixed effects as we do in our benchmark specification, then the instrument picks up variation around the country mean. Intuitively, holding that mean fixed, when forecasters outside of the domestic country of bank b are more optimistic than this average, then naturally the excluded forecasters (like bank b) are likely to be below the average forecast for that country, at that time period. In accordance with this intuition, untabulated first stage regression results with country and time fixed effects included separately deliver a positive coefficient on $F(Y10)_{nonbank}$.

2.1.4 Domestic bond holdings. While our primary focus is on the link between information heterogeneity and foreign portfolio holdings, in Internet Appendix Table A9 we present regressions estimates where we also include domestic portfolio holdings on the left-hand side of Equation (3). The results are similar, even quantitatively. Interestingly, the coefficient for a dummy for domestic assets is large, positive, and significant, which suggests that information frictions cannot fully explain the home bias in our data. Thus, at least in our sample of sophisticated investors, information heterogeneity plays a crucial role in generating heterogeneity across individual foreign portfolio holdings, but cannot explain the underweighting of foreign assets as a whole relative to the home asset.

2.2 Sparseness in information and portfolios

Lastly, we explore the relationship between the extensive margin of information and portfolio holdings. This analysis aims to leverage the full extent of our data, as it allows us to include bank-country observations where we have portfolio data, but not a Consensus forecast.

Specifically, we test whether reporting a forecast for the 10-year yield of a foreign country predicts investments in that country's sovereign debt by estimating:

$$Share_{bct} = \mu_{bt} + \gamma_{ct} + \beta_D \ln(Dist)_{bct} + \beta_F \mathbb{1}(ForeignFcst_{bct}) + \varepsilon_{bct} \qquad (6)$$

where $\mathbb{1}(ForeignFcst_{bct})$ is a dummy variable that equals one if bank b makes a yield forecast about country c at time t and zero otherwise; μ_{bt} and γ_{ct} represent bank-time and destination country-time fixed effects. As before, we also control for the effect of gravity. In this regression we can make use of all bank-country-quarter observations, not just the ones for which a specific forecast is reported, as was the case in Table 3. However, the trade-off is that we cannot estimate the specific effect of information precision or the elasticity of portfolios to point forecasts. Still, our basic hypothesis is that observing a forecast is indicative of the bank having a macro research team covering that particular country, and thus the regression can help us evaluate if information production is indeed associated with portfolio holdings in the broadest sample of our data.

Panel A of Table 6 presents the results. In line with previous results, we again find a significant gravity relationship. In the most saturated specification that includes both bank-time and country-time fixed effects (column 4), the coefficient of distance implies that doubling the distance between home and foreign countries is associated with a 3.6 percentage points decrease in the portfolio share of that sovereign, which is again significant compared to the average portfolio share of an individual foreign debt holding of 2.28% (Table 1, panel A).

In columns 5 to 8, we include the indicator variable $\mathbb{1}(ForeignFcst_{bct})$, and find a positive coefficient that is stable across specifications, suggesting a robust relationship between information production and investment. The results are

Table 6
Sparseness in foreign portfolios and forecasts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Dependent variable $Share_{bct}$								
log(dist)	−3.385***	−3.575***	−3.565***	−3.565***	−3.170***	−3.388***	−3.285***	−3.261***
	(0.590)	(0.566)	(0.688)	(0.702)	(0.586)	(0.572)	(0.699)	(0.714)
$\mathbb{1}$(ForeignFcst)					2.106**	1.766*	1.708*	1.854*
					(0.950)	(0.952)	(0.882)	(0.957)
Observations	3,080	3,080	3,080	3,080	3,080	3,080	3,080	3,080
R^2	.100	.139	.306	.310	.110	.146	.312	.317
B. Dependent variable $100\times\mathbb{1}(Share_{bct})$								
log(dist)	−22.51***	−21.76***	−24.10***	−24.10***	−19.60***	−19.89***	−21.42***	−21.39***
	(3.07)	(2.35)	(2.73)	(2.79)	(2.99)	(2.23)	(2.54)	(2.60)
$\mathbb{1}$(ForeignFcst)					28.53***	17.62***	16.32***	16.51***
					(4.31)	(3.96)	(3.93)	(4.23)
Time FE	No	Yes	–	–	No	Yes	–	–
Bank FE	No	Yes	Yes	–	No	Yes	Yes	–
Country-time FE	No	No	Yes	Yes	No	No	Yes	Yes
Bank-time FE	No	No	No	Yes	No	No	No	Yes
Observations	3,080	3,080	3,080	3,080	3,080	3,080	3,080	3,080
R^2	.081	.365	.514	.529	.118	.378	.524	.538

This table provides the estimates for Equation (6). The dependent variable is the share of foreign EEA country c sovereign bonds in bank b sovereign portfolio of long-term debt (> 5 years residual maturity) in panel A and a dummy equal to one if bank b holds a positive amount (in excess of one basis point of the total sovereign portfolio) of long-term sovereign bonds of EEA country c in panel B. log(dist) is the natural logarithm of the bilateral distances between country c and the home country of bank b, weighted by population of the major cities (Mayer and Zignago 2011). $\mathbb{1}$(ForeignFcst) is a dummy equal to one if bank b makes a 10-year yield forecast for country c in year t and zero otherwise. Standard errors are reported in parentheses and are clustered at the bank-country level. *p <.1; **p <.05; ***p <.01.

both statistically and economically significant, and imply that making a forecast about a given country is associated with roughly a 2 percentage points larger portfolio share for that country's sovereign debt.

Finally, since bank portfolios are also not necessarily invested in all foreign countries, we test whether information production is associated with the bank taking a position in the country's sovereign debt at all (i.e., the extensive margin of investment). To do so, we modify Equation (6) by replacing the left-hand side with the indicator variable $\mathbb{1}(Share_{b,c,t})$, that is equal to one if at time t bank b holds a positive amount of country c's sovereign debt (in excess of one basis point of the total sovereign portfolio), and zero otherwise.[21] The estimates are presented in panel B of Table 6, and show a strong information effect: the coefficient for $\mathbb{1}(ForeignFcst_{bct})$ is positive and significant across specifications, indicating that making a forecast for a foreign country is associated with 16 to 28 percentage points higher probability of investing in that country's sovereign (columns 5 to 8).

Thus, we broadly confirm the basic findings of our benchmark regressions in this much bigger sample. Nevertheless, the simple indicator function of whether a bank produces a forecast is at best a coarse measure of information production, and, hence, it is not surprising that the effect of gravity remains significant after including the binary information variable.

3. Model

Our empirical findings not only speak to the important role played by information in determining international portfolios but also shed light on the empirically relevant primitives of the structure of information production, which is an important open question in the theoretical literature.

Specifically, Van Nieuwerburgh and Veldkamp (2010) show that different commonly used assumptions on the information production technology can lead to vastly different equilibrium portfolio allocations, from very concentrated to perfectly diversified portfolios. With this in mind, Van Nieuwerburgh and Veldkamp (2009) exploit the presence of home bias in portfolios to argue for a specific formulation of the basic framework that leads to a full specialization in information production. In that case, agents choose to have forecasting expertise only over their home asset, in turn generating significant home bias in portfolios in equilibrium.

Our empirical findings help us further discipline the information structure and thus the underlying primitives of the model. Specifically, full specialization in information is inconsistent with our finding of significant heterogeneity in the ability of investors to forecast different foreign assets (Figure 2), and that

21 Further details on the sparseness of portfolios and different margins of international diversification are found in Internet Appendix A.4.

such heterogeneity explains international portfolios both across (Table 3) and within investors (Table 4).

In this section, we show the standard framework can be made to fit the richer dimensions of the data that we document, by introducing just one intuitive change to the baseline Van Nieuwerburgh and Veldkamp (2009) model. Namely, we introduce a two-factor asset payoff structure where information can be produced about only one of the factors. This introduces a measure of "unlearnable" uncertainty in asset payoffs, which we show weakens the feedback between portfolio and information choice that underpins the standard result of full specialization in information production. In this way, we obtain an equilibrium model where investors optimally produce costly information for a subset of all assets, which gives rise to partial information specialization and heterogeneity in forecasting expertise. Overall, we show that unlearnable uncertainty is not just an intuitive, but also a necessary component for information-based models of international portfolio choice.

3.1 Model description

The model has three periods. In period 1, agents choose their information production strategy, which describes how much information to produce for each asset available for trading. In period 2, the resultant informative signals realize, agents update their priors, and form optimal portfolios; in period 3 agents consume the resultant returns on their portfolios.

There are N countries, with a continuum of agents of mass $\frac{1}{N}$ in each, N risky assets, one associated with each country, and a risk-free savings technology with an exogenous rate of return normalized to one. In period 2 agent i in country j faces the budget constraint

$$W_{2j}^{(i)} = \sum_{k=1}^{N} P_k x_{jk}^{(i)} + b_j^{(i)},$$

where P_k is the price of the risky asset of country k, $x_{jk}^{(i)}$ are the portfolio holdings of risky assets, $b_j^{(i)}$ are risk-free savings and $W_{2j}^{(i)}$ is the total investable wealth of the agent. To reduce clutter, we suppress the i index (denoting agents) if there is no chance of confusion.

The assets yield stochastic payoffs, which we collect in the N-by-1 vector $\mathbf{d}$. We assume that the payoffs are driven by two independent factors, so that $\mathbf{d} = \mathbf{f} + \mathbf{e}$, where both $\mathbf{f}$ and $\mathbf{e}$ are independent Gaussian vectors: $\mathbf{f} \sim N(\mu_f, \Sigma_f)$ and $\mathbf{e} \sim N(\mu_e, \Sigma_e)$.[22] The key difference between the two factors is that agents

[22] We follow the convention of assuming Gaussian payoffs to be consistent with the previous literature on asset pricing with endogenous information acquisition. Moreover, in our quarterly data set of long-term sovereign bonds we cannot reject the hypothesis that the distribution of bond returns is indeed Normal, likely reflecting the fact that yields are relatively far from the effective lower bound. Nevertheless, our model can be modified with non-Normal payoffs along the lines of Albagli et al. (2021) and Breon-Drish (2015), for example, who show

are able to produce information about $\mathbf{f}$, but not about $\mathbf{e}$. We refer to $\mathbf{f}$ as the "fundamental" factor, which is meant to model fluctuations in returns related to forecastable fundamentals. The factor $\mathbf{e}$ captures fluctuations in returns that are unrelated to any information that is potentially available at the time of trading. That could be interpreted as unforecastable fluctuations in future fundamentals, but also random and unforecastable shocks to sentiments or risk premiums more broadly. For tractability, we assume that the variance matrices Σ_f and Σ_e are diagonal, but introducing a factor structure in the cross-section of returns does not alter the results. We also normalize $\mu_e = 0$, as the unconditional mean of this unlearnable component makes no difference to the analysis.

In forecasting returns, investors observe the equilibrium prices $\mathbf{p}$ and optimally extract any information about payoffs from those prices. This information is public and free. Moreover, the investors can produce additional information in the form of unbiased signals about the realization of the different countries' fundamental factors f_k:

$$\eta_{jk}^{(i)} = f_k + u_{jk}^{(i)},$$

where $u_{jk}^{(i)} \overset{iid}{\sim} N(0, \sigma_{u_{jk}}^{(i)2})$ is an idiosyncratic error term. The precision of the signals, $\frac{1}{\sigma_{u_{jk}}^{(i)2}}$, is chosen optimally, subject to an increasing and convex cost, $C(K)$, of the total amount of information, K, encoded in their chosen signals. We measure information flow in terms of reduction in uncertainty about $\mathbf{f}$ conditional on observing the vector of signals $\boldsymbol{\eta}_j^{(i)} = [\eta_{j1}, \ldots, \eta_{jN}]'$, where uncertainty is measured in entropy units. Formally,

$$K = H(\mathbf{f}|\mathcal{I}^p) - H(\mathbf{f}|\mathcal{I}^p, \boldsymbol{\eta}_j^{(i)}),$$

where $H(X)$ is the entropy of random variable X, $H(X|Y)$ is the entropy of X conditional on knowing Y, and $\mathcal{I}^p$ is the public information set, which contains both the common priors and the freely observed equilibrium vector of asset prices $\mathbf{p}$.[23] Thus, K measures the total amount of information about the vector of unknown fundamentals $\mathbf{f}$ contained in the vector of private signals, $\boldsymbol{\eta}_j^{(i)}$, over and above the publicly available information. Lastly, as is standard in the home bias literature, we endow agents with a small exogenous information advantage over their home asset, by assuming that the first ε units of entropy about the home asset are free.[24] This is not necessary for our key results on the

that, taking information structure as exogenous, under certain conditions the equilibrium prices are still available in closed form. In particular, Breon-Drish (2015) shows that the price functions may also remain monotonic in signal precision, which is a key feature of our model too, suggesting that the optimal information problem should yield similar qualitative implications.

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

[24] It is convenient to introduce the home information advantage exogenously, but it can be endogenized as in Nieuwerburgh and Veldkamp (2006) and Valchev (2017) by modeling domestic nontradable income (e.g., commercial loans).

equilibrium information and portfolio heterogeneity across *foreign* assets, and is only used to generate some home bias.

We view this information structure as a simple, yet realistic way of modeling the costs banks incur in researching and producing information that would be useful in forecasting future bond yields. We similarly think that any information that can be inferred by the current level of asset prices is fully internalized by sophisticated investors, and, hence, we model that information as free. However, to create further value, financial institutions also dig deeper and hire teams of analysts and "quants," to build more comprehensive and insightful forecasting models. Building and maintaining such forecasting teams is expensive, and banks need to choose optimally how to allocate these costly resources in terms of producing information about all N countries, which leads to dispersion in views and forecasting expertise across investors as we describe next.

3.2 Period 2: Portfolio choice

In period 2, agents update their priors about the asset payoffs with the realizations of the signals $\boldsymbol{\eta}_j^{(i)}$, and then form optimal portfolios that maximize expected utility over their terminal wealth. We solve the model by backward induction, starting with the optimal portfolio choice in period 2, and then solving for the optimal information choice in period 1.

Investor i in country j picks the portfolio composition that maximizes expected utility:

$$\max_{\boldsymbol{x}_j^{(i)\prime}} E\left[\frac{(W_{3j}^{(i)})^{1-\gamma}}{1-\gamma}|\mathcal{I}_j^{(i)}\right] \text{ subject to}$$

$$W_{3j}^{(i)} = W_{2j}^{(i)}(1+\frac{1}{W_{2j}^{(i)}}\boldsymbol{x}_j^{(i)\prime}(\mathbf{d}-\mathbf{p})),$$

where $W_{2j}^{(i)} = W_1 - C(K_j^{(i)})$ is the net invested wealth of the agent, equal to the common initial endowment (W_1) minus the information production costs ($C(K_j^{(i)})$), and $\mathcal{I}_j^{(i)} = \{\mathcal{I}_p, \boldsymbol{\eta}_j^{(i)}\}$ is the information set of investor i in country j. We follow Campbell and Viceira (2001) and use a second-order Taylor expansion to solve the model. The technical details of the solution method are standard and are described in detail in Internet Appendix B.

The optimal portfolio holdings of asset k for agent i are given by:

$$\frac{x_{jk}^{(i)}}{W_{2j}^{(i)}} = \frac{E(d_k|\mathcal{I}_j^{(i)}) - p_k}{\gamma(\hat{\sigma}_{jk})^2} \tag{7}$$

where $\hat{\sigma}_{jk}^2$ is the kth diagonal element of the posterior variance matrix of asset payoffs $\hat{\Sigma}_j^d = \mathrm{Var}(\mathbf{d}|\mathcal{I}_j^{(i)})$. Thus, the agent invests more heavily in assets she expects to do better (high $E(d_k|\mathcal{I}_j^{(i)})$, and invests less in more uncertain assets that have a higher posterior variance. And, of course, there is a key interaction:

the higher the precision of beliefs, the higher is the elasticity of portfolio holdings to expected payoffs. The crucial question is, of course, whether the (endogenous) equilibrium information structure is indeed such that the posterior variance of beliefs varies across foreign assets in the way we found it does in the data.

In addition, there are also noise traders that trade the N assets for reasons orthogonal to the asset payoffs. From a technical perspective, they are needed to ensure more shocks than asset prices so that private information does not unravel in equilibrium. In practice, a substantial amount of bonds is held for liquidity and hedging purposes and, to the extent to which those motives are unrelated to yield fluctuations (i.e., the fundamentals), they are modeled by the noise trading shocks. Market clearing thus requires,

$$\sum_{j=1}^{n}\int x_{jk}^{(i)}di=z_k \tag{8}$$

and noise trading effectively functions as stochastic asset supply, where $z_k \overset{iid}{\sim} N(\mu_{zk},\sigma_{zk}^2)$.

We assume that all assets are ex ante identical. Thus, we set the mean of the supply of each asset $\mu_{zk}=\frac{1}{N}$ for all k, and assume that the ex ante expected payoff vector μ_f has all the same elements, and that the variances of the payoff factors are given by the identity matrix scaled by scalars σ_f^2 and σ_e^2 that capture the common variance of the fundamental and unlearnable components, respectively (i.e., $\Sigma_f=\sigma_f^2 I_N$ and $\Sigma_e=\sigma_e^2 I_N$).

Lastly, we focus on symmetric equilibria, where all agents in a country make the same information choices, and guess and verify that the equilibrium price is of the linear form,

$$p_k=\bar{\lambda}_k+\lambda_{fk}f_k+\lambda_{zk}z_k,$$

following Admati (1985) (see Internet Appendix B for the details).

3.3 Period 1: Information choice

Investors make their information production choices at time 1, before asset markets open. We solve the information choice problem in two steps. First, we solve for the optimal allocation of information across the N different country fundamentals, that is, the set of $\{\kappa_{jk}\}$, given a total level of information production K_j. Second, given this optimal information *allocation* strategy, we solve for the optimal level of total information production K_j^*.

3.3.1 Optimal allocation of information production. Our main result here is that, because of unlearnable uncertainty ($\sigma_e^2>0$), the problem of allocating information to individual assets is convex at low levels of information, but becomes concave after the allocated information exceeds an asset-specific threshold level $\bar{\kappa}_{jk}>0$.

In other words, at the onset of learning about an asset k an investor experiences increasing returns to producing additional information about that asset, but as information accumulates, the marginal benefit of additional information starts declining. As a result, learning proceeds in a cascading "water-filling" way. Specifically, there exist thresholds of total information acquired $\{\bar{K}_{j1},\ldots,\bar{K}_{j,N-1}\}$ such that if $K_j < \bar{K}_{j1}$ the agent allocates all of his information production to just one asset: that is, $\kappa_{jk} = K_j$, and $\kappa_{jk'} = 0$ for all $k' \neq k$. If $K_j \in [\bar{K}_{j1}, \bar{K}_{j2})$, then the agent splits the acquired information across two assets, and so on. Thus, for any level of total information production $K_j < \bar{K}_{j,N-1}$, an agent only produces additional information for a subset of the assets, and only if $K_j > \bar{K}_{j,N-1}$ he finds it optimal to produce additional information for all assets. Proposition 1 formalizes this result.

Proposition 1. Extensive margin in information choice. If $\sigma_e^2 \in (0, \bar{\sigma}_e^2)$, where $\bar{\sigma}_e^2$ is defined in Internet Appendix C, there exist positive constants $\bar{K}_{j1} < \bar{K}_{j2} < \cdots < \bar{K}_{j,N-1}$ such that if

- $K_j \leq \bar{K}_{j1}$, an agent produces additional information only for their domestic asset: $\kappa_{jj} = K_j$, $\kappa_{jk} = 0$ for all other k
- $K_j \in (\bar{K}_{j,L-1}, \bar{K}_{jL}]$, an agent produces additional information for $L > 1$ countries: $\kappa_{jk} > 0$ for L different countries with $\sum_{k=1}^{L} \kappa_{jk} = K_j$, and $\kappa_{jk'} = 0$ for all other k'.
- $K_j > \bar{K}_{j,N-1}$, an agent producess information for all assets: $\kappa_{jk} > 0$ $\forall k$, $\sum_k \kappa_{jk} = K_j$

Proof. The intuition is sketched above, and the details are presented in Internet Appendix C. ■

The increasing returns to information come about due to a feedback between information and portfolio choices, as also present in the benchmark model of Van Nieuwerburgh and Veldkamp (2009). Allocating more information production to a given asset k reduces the posterior uncertainty about its return, and thus the investor optimally increases her holdings of that asset. As the holdings increase, utility becomes more sensitive to fluctuations in the return of asset k and the value of information about asset k increases.

In our model, however, there is also an additional effect. Increasing the portfolio share of asset k exposes the investor to increasing amounts of unlearnable risk e_k. This moderates the incentive to increase the portfolio share and weakens the feedback loop described above. The key result is that the feedback effect is strong, and thus resulting in increasing returns to information, when the investor has not yet acquired much information about f_k and the posterior variance of the learnable fundamental is high relative to that of e_k. As increased information production shrinks the posterior variance of f_k, an increasingly larger proportion of the residual uncertainty in the return on asset k is driven by the unlearnable risk. Thus, further information acquisition has progressively smaller ability to shrink the remaining uncertainty, weakening the feedback effect with the portfolio holding. Eventually, as the share of learnable uncertainty shrinks, this leads to a decreasing marginal benefit to information about asset k.

This basic feature of the model generates an extensive margin in information production. Because of the initially increasing returns, an agent does not necessarily produce additional information about all assets. The initial increasing returns to information, combined with the small ex ante advantage in home information, ensures that every investor will at first produce information about their home asset. However, agents do not necessarily stop there. If the chosen amount of total information production ability, K_j, is high enough, spending it all on home information becomes suboptimal because of the eventually decreasing returns to further home information. Hence, investors with a sufficiently high K_j allocate their information production to more than one asset. New assets are added in a stepwise fashion, however, because each enjoys increasing returns to information initially.

We call this "partial information specialization" as investors generally specialize in producing information for a *nonsingleton* subset of countries. This generates significant heterogeneity in the precision of beliefs across different assets, as is also true in the data (Section 1.2).

Naturally, in the limiting case of $\sigma_e^2 \to 0$, we are back to the original Van Nieuwerburgh and Veldkamp (2009) result that investors fully specialize in the production of information about their domestic asset only. In this case, the only information heterogeneity is between home and foreign assets as a whole, but not within foreign assets themselves.

Corollary 1. As $\sigma_e^2 \to 0$, then $\bar{K}_{j,1} \to \infty$ and agents only produce home information.

Lastly, we note that if unlearnable uncertainty is too large (i.e., σ_e^2 larger than a threshold value $\bar{\sigma}_e^2$ we derive in Internet Appendix C), then information acquisition always displays decreasing returns and thus agents produce information about all assets (i.e., $\bar{K}_{j,N-1} = 0$). In such a case, the fraction of overall uncertainty that can be reduced through information acquisition is so small that the feedback between information and portfolio holdings is always too weak to generate increasing returns to information. This is not a particularly interesting or empirically relevant case, and, hence, for the results below we will assume that $\sigma_e^2 < \bar{\sigma}_e^2$.

Given the optimal information allocation strategy we characterized above, we can then solve for the total amount of information agents produce, K_j. This choice does not play a crucial role in the key result of heterogeneity in information precision across different foreign assets (i.e., partial specialization), and, hence, we leave the derivation of the optimal K_j to Internet Appendix B.2.2.

3.4 Key implications of the model

Most importantly, the model is able to match the novel empirical regularities we documented earlier. First, thanks to unlearnable uncertainty, that is, $\sigma_e^2 > 0$, and the resultant partial information specialization derived in Proposition 1, the model generates cross-sectional heterogeneity in the precision of beliefs about different foreign assets at the investor level. That is, $\widehat{\sigma}_{jk}^2$ varies at the investor-asset level for $j \neq k$, as investors produce information about some, but not all, foreign assets. In turn, through Equation (7), this generates heterogeneity in portfolio weights at the investor-asset level, such that the portfolio share of asset k and the elasticity of that portfolio share to the point forecast $E(d_k|\mathcal{I}_j)$ are both increasing in the specific investor-asset level of information precision $\frac{1}{\widehat{\sigma}_{jk}^2}$, just as we found to be true in the data.

Proposition 2. Equilibrium portfolio and information heterogeneity. Assuming that information production is not too big or too small ($K_j \in (\bar{K}_{j,1}, \bar{K}_{j,N-1})$):

1. **Concentration of foreign holdings:** Any foreign asset k that an investor produces information about ($\kappa_{jk} > 0$) is overweighed relative to the market portfolio, while any foreign asset k' the investor does not produce information for ($\kappa_{jk'} = 0$) is underweighted:
$$E\left(\frac{x_{jk}}{W_{2j}}\right) - \mu_z > 0 > E\left(\frac{x_{jk'}}{W_{2j}}\right) - \mu_z$$
2. **Information precision:** For such asset pairs k and k', where $\kappa_{jk} > 0$ and $\kappa_{jk'} = 0$
$$\widehat{\sigma}_{jk}^2 < \widehat{\sigma}_{jk'}^2 \Rightarrow E\left(\frac{x_{jk}}{W_{2j}}\right) > E\left(\frac{x_{jk'}}{W_{2j}}\right)$$
3. **Portfolio sensitivity to expectations:** Portfolio shares of the assets that investors produce information about are more sensitive to point forecasts of returns:
$$\frac{\partial x_{jk}/W_{2j}}{\partial E(d_k|\mathcal{I}_j)} > \frac{\partial x_{jk'}/W_{2j}}{\partial E(d_{k'}|\mathcal{I}_j)} \iff \kappa_{jk} > 0 \text{ and } \kappa_{jk'} = 0$$

Proof. The intuition is sketched in the text, and details are in Internet Appendix C. ■

The proposition leverages the key feature of optimal information production that investors produce information about some, but not all, foreign countries, as follows.[25]

[25] The assumption that $K_j \in (\bar{K}_{j,1}, \bar{K}_{j,N-1})$ simply excludes extreme parametrizations where investors either do not produce information about any foreign assets ($K_j < \bar{K}_{j,1}$) or produce information about all foreign assets ($K_j > \bar{K}_{j,N-1}$).

First, the equilibrium portfolio holdings align with the first stylized fact we documented in Section 1, namely that the portfolio allocations across foreign assets deviate from CAPM. Intuitively, in our model investors produce information about more than just their home asset, which creates information asymmetry not only between home and foreign assets as a whole (as in the benchmark Van Nieuwerburgh and Veldkamp (2009) model) but also within the set of foreign assets. Specifically, investors have higher precision of beliefs for the subset of foreign assets for which they produce extra information, and thus will overweight those assets in their portfolio as per the optimal portfolio holdings expression in Equation (7). On the other hand, for the foreign assets for which an investor does not produce additional information, the precision of beliefs is relatively low, and, hence, she ends up underweighting those assets in her portfolio. Put differently, in equilibrium the supply of any given asset is concentrated in the hands of the investors that are best informed about its fundamentals, leading to deviations from CAPM that also align with these endogenous, comparative advantages in information.

The fact that the systematic deviation from CAPM are driven by investor-asset-specific differences in information precision directly implies the second result of the proposition, which speaks to two different empirical results we have documented. First, it shows that investors have heterogeneous ability to forecast different asset payoffs. Since $\widehat{\sigma}_{jk}^2 = E(d_k - E(d_k \big| \mathcal{I}_j^{(i)}))^2$ is the model's direct counterpart to the squared forecast error variance that we compute in Section 1.2, the first part of the result shows that agents in our model indeed have heterogeneous ability to forecast the returns of different (foreign) assets. Second, this heterogeneity in forecast precision allows the model to capture the negative relationship between forecast error variance and portfolio shares across investor-asset pairs that we estimate in the data (i.e., this delivers a $\beta_1 < 0$ in regression (3)).

Lastly, the third result of Proposition 2 formalizes the link between information production and the sensitivity of portfolio holdings to the point forecast of asset returns, which captures our most important empirical result, namely, the positive interaction between the a forecast and its precision ($\beta_3 > 0$ in regression (3)). This key empirical result shows that investors are indeed exploiting their comparative advantage in information that supports the key tenet of information-based models. In our model, the comparative advantage in information arise *endogenously*, due to the interaction of costly information and unlearnable uncertainty, which leads to partial information specialization, heterogeneity in forecast precision, and thus increased sensitivity to the specific point forecast (Equation (7)).

While costly information being necessary for these results is perhaps obvious, the role played by unlearnable uncertainty is more subtle, but similarly crucial. Specifically, in the limiting case of $\sigma_e^2 \to 0$, there is no heterogeneity in information production across *foreign* assets in equilibrium, and, hence, all three results of Proposition 2 disappear.

Corollary 2. As unlearnable uncertainty becomes negligible, $\sigma_e^2 \to 0$,

- The relative CAPM deviations across foreign assets are equalized: $E(\frac{x_{jk}}{W_{2j}}) - \mu_z \to 0$
- Equal information about all foreign assets: $\widehat{\sigma}_{jk}^2 - \widehat{\sigma}_{jk'}^2 \to 0 \Rightarrow E(\frac{x_{jk}}{W_{2j}}) - E(\frac{x_{jk'}}{W_{2'}}) \to 0$
- Sensitivity to expectations is the same for all foreign assets: $\frac{\partial x_{jk}/W_{2j}}{\partial E(d_k|\mathcal{I}_j)} - \frac{\partial x_{jk'}/W_{2j}}{\partial E(d_{k'}|\mathcal{I}_j)} \to 0$

Without unlearnable uncertainty, the model implies that the coefficient for the average precision of bank forecasts (β_2 in regression (3)) should be completely absorbed by the fixed effects and that the key coefficient for the interaction between point forecasts and the forecast precision should be zero ($\beta_3 = 0$). Both implications are clearly counterfactual. Hence, the model suggests that unlearnable uncertainty is a necessary ingredient for information models that want to match the data on international portfolio holdings.

In closing, we note that the unique partial information specialization feature of our model implies that the choice of *how many* and *which* foreign countries each investor decides to learn about is endogenous. This novel feature of our model implies that *bilateral* portfolio flows experience volatility if investors decide to stop learning about an asset they had been learning about in the past. Such a shrinking of the set of assets being learned leads to what can look like capital flows “retrenchment” and “fickleness,” as documented by Caballero and Simsek (2020). Our model thus provides a novel interpretation, based on endogenous information reallocation, of these other empirical regularities. We explore these implications in more detail and confirm them empirically in Internet Appendix D.

4. Conclusion

In this paper, we document that information heterogeneity is crucial to explain the heterogeneity in the allocation of international investments. Our results are consistent with the implications of information-based models of portfolio choice. In particular, it shows that individual portfolio holdings are more sensitive to point forecasts that are objectively more accurate. This means that investors internalize where their comparative advantage in information lies and form portfolios accordingly. Finally, we provide a model with a two-tiered information structure featuring unlearnable uncertainty in asset payoffs that eventually generates decreasing returns to information.

Our results have a number of broad implications for policy and future research. We document that the Consensus Economics survey indeed contains valuable information that can be used to understand observed portfolio structures, help policymakers monitor and predict capital flows based on the relative optimism of banks located in different countries, and also help investors generate excess returns. Second, our model shows that unlearnable uncertainty is a necessary ingredient to match the data, suggesting that it should become an integral part of the benchmark information friction framework.

References

Admati, A. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53:629–58.

Ahearne, A., W. Griever, and F. Warnock. 2004. Information costs and home bias: an analysis of us holdings of foreign equities. *Journal of International Economics* 62:313–36.

Albagli, E., C. Hellwig, and A. Tsyvinski. 2021. Dispersed information and asset prices. Working Paper, Toulouse School of Economics.

Albuquerque, R., G. H. Bauer, and M. Schneider. 2009. Global private information in international equity markets. *Journal of Financial Economics* 94:18–46.

Bohn, H., and L. L. Tesar. 1996. Us equity investment in foreign markets: portfolio rebalancing or return chasing? *The American Economic Review* 86:77–81.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer. 2020. Overreaction in macroeconomic expectations. *American Economic Review* 110:2748–82.

Breon-Drish, B. 2015. On existence and uniqueness of equilibrium in a class of noisy rational expectations models. *The Review of Economic Studies* 82:868–921.

Caballero, R. J., and A. Simsek. 2020. A model of fickle capital flows and retrenchment. *Journal of Political Economy* 128:2288–328.

Campbell, J. Y., L. E. Calvet, and P. Sodini. 2007. Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy* 115:707–47.

Campbell, J. Y., and L. M. Viceira. 2001. Who should buy long-term bonds? *American Economic Review* 91:99–127.

Chan, K., V. Covrig, and L. Ng. 2005. What determines the domestic bias and foreign bias? evidence from mutual fund equity allocations worldwide. *The Journal of Finance* 60:1495–534.

Chari, V., A. Dovis, and P. Kehoe. 2020. On the optimality of financial repression. *Journal of Political Economy* 128:710–39.

Coeurdacier, N., and H. Rey. 2012. Home bias in open economy financial macroeconomics. *Journal of Economic Literature* 51:63–115.

Coibion, O., and Y. Gorodnichenko. 2015. Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review* 105:2644–78.

Coval, J. D., and T. J. Moskowitz. 2001. The geography of investment: Informed trading and asset prices. *Journal of Political Economy* 109:811–41.

Curcuru, S. E., C. P. Thomas, F. E. Warnock, and J. Wongswan. 2011. Us international equity investment and past and prospective returns. *American Economic Review* 101:3440–55.

De Marco, F., and M. Macchiavelli. 2015. The political origin of home bias: The case of europe. Working Paper, Federal Reserve Board.

Dziuda, W., and J. Mondria. 2012. Asymmetric information, portfolio managers, and home bias. *Review of Financial Studies* 25:2109–54.

ESRB. 2015. Report on the regulatory treatment of sovereign exposures. Report, European Systemic Risk Board, https://www.esrb.europa.eu/pub/pdf/reports/esrbreport...s032015.en.pdf.

Frankel, J., and C. M. Engel. 1984. Do asset-demand functions optimize over the mean and variance of real returns? a six-currency test. *Journal of International Economics* 17:309–23.

French, K., and J. Poterba. 1991. Investor diversification and international equity markets. *American Economic Review* 81:222–6.

Froot, K. A., P. G. OâŁ™connell, and M. S. Seasholes. 2001. The portfolio flows of international investors. *Journal of Financial Economics* 59:151–93.

Gelos, R. G., and S.-J. Wei. 2005. Transparency and international portfolio holdings. *The Journal of Finance* 60:2987–3020.

Gennaioli, N., Y. Ma, and A. Shleifer. 2016. Expectations and investment. *NBER Macroeconomics Annual* 30:379–431.

Giacoletti, M., K. Laursen, and K. J. Singleton. 2021. Learning from disagreement in the u.s. treasury bond market. *The Journal of Finance* 76:395–441.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus. 2021. Five facts about beliefs and portfolios. *American Economic Review* 111:1481–522.

Grinblatt, M., and M. Keloharju. 2001. How distance, language, and culture influence stockholdings and trades. *The Journal of Finance* 56:1053–73.

Guiso, L., and T. Jappelli. 2008. Financial literacy and portfolio diversification. Working Paper, Einaudi Institute for Economics and Finance.

Guiso, L., P. Sapienza, and L. Zingales. 2009. Cultural biases in economic exchange? *The Quarterly Journal of Economics* 124:1095–131.

Hau, H., and H. Rey. 2008. Home bias at the fund level. *The American Economic Review P&P* 98:333–8.

Huberman, G. 2001. Familiarity breeds investment. *Review of Financial Studies* 14:659–80.

Ivković, Z., and S. Weisbenner. 2005. Local does as local is: Information content of the geography of individual investors' common stock investments. *The Journal of Finance* 60:267–306.

Koijen, R. S., and M. Yogo. 2020. Exchange rates and asset prices in a global demand system. Working Paper, National Bureau of Economic Research.

Leuz, C., K. V. Lins, and F. E. Warnock. 2009. Do foreigners invest less in poorly governed firms? *The Review of Financial Studies* 22:3245–85.

Mayer, T., and S. Zignago. 2011. Notes on cepii's distances measures : the geodist database. Working Paper, CEPII.

Merton, R. C. 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics* 247–57.

Mondria, J. 2010. Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory* 145:1837–64.

Nechio, F. 2014. Foreign stock holdings: The role of information. Working Paper, San Francisco Federal Reserve.

Nieuwerburgh, S., and L. Veldkamp. 2006. Inside information and the own company stock puzzle. *Journal of the European Economic Association* 4:623–33.

Ongena, S., A. Popov, and N. Van Horen. 2019. The invisible hand of the government: Moral suasion during the european sovereign debt crisis. *AEJ: Macroeconomics* 11:346–79.

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

Portes, R., and H. Rey. 2005. The determinants of cross-border equity flows. *Journal of international Economics* 65:269–96.

Stavrakeva, V., and J. Tang. 2020. Deviations from fire and exchange rates: A ge theory of supply and demand. Working Paper.

Valchev, R. 2017. Dynamic information acquisition and portfolio bias. Working Paper, Boston College.

Van Nieuwerburgh, S., and L. Veldkamp. 2009. Information immobility and the home bias puzzle. *The Journal of Finance* 64:1187–215.

———. 2010. Information acquisition and under-diversification. *Review of Economic Studies* 77:779–805.

Wang, C. 2020. Under-and over-reaction in yield curve expectations. Working Paper, University of Notre Dame.