The Geography of Beliefs*

Harjoat S. Bhamra   Raman Uppal   Johan Walden

June 11, 2020

Abstract

Empirical evidence shows that beliefs of households deviate from rational expectations. We develop a model where a household's beliefs about stock returns are an *endogenous* outcome of its location in a bipartite network of households and firms. To test this model, we establish the relation between households' beliefs and their portfolio choices and exploit Finnish data for 125 stocks and the portfolio holdings of 405,628 households. We find that household-firm distance in the network has a statistically and economically significant effect on household beliefs about firm-level stock returns implying that geography has a strong effect on beliefs. We calculate the reduction in household welfare resulting from the deviation of beliefs from rational expectations and show how this varies depending on the location of households in the network.

Keywords: household finance, behavioral finance, networks.
JEL: D84, E03, G02, G11.

*Harjoat Bhamra is affiliated with Imperial College Business School and CEPR; Email: h.bhamra@imperial.ac.uk. Raman Uppal is affiliated with Edhec Business School and CEPR; Email: raman.uppal@edhec.edu. Johan Walden is affiliated with University of Lausanne, Swiss Finance Institute, and UC Berkeley, Haas School of Business; Email: johan.walden@unil.ch. We thank David Sraer for valuable suggestions, and Nicholas Bel and Zhimin Chen for research assistance. We are grateful for comments from Hossein Asgharian, Claes Bäckman, Olga Balakina, Adrian Buss, Veronika Czellar, Winifred Huang, Abraham Lioui, Massimo Massa, David Newton, Kim Peijnenburg, Joël Peress, Argyris Tsiaras, Ania Zalewska, and seminar participants at Bocconi University, INSEAD, Lund University, University of Bath, and University of Cambridge.

1 Introduction and Motivation

Rational expectations is a mathematically convenient fiction beloved of economists. However, "the rational expectation hypothesis is strongly rejected" (Landier, Ma, and Thesmar, 2017). This finding necessitates taking a stance on how to deviate from rational expectations, while still imposing rigor on the manner in which beliefs are formed. In this paper, we develop a theoretical framework to demonstrate how one can deviate from rational expectations in a disciplined fashion, and then establish empirically using Finnish data on the location of firms and the portfolio holdings of households that geography has a strong effect on beliefs about stock returns.

In the framework we develop, households' beliefs are derived *endogenously* based on their location relative to firms within a bipartite network.[1] Networks are ubiquitous in modern economies: "Networks determine our information, influence our opinions, and shape our political attitudes" (Acemoglu and Ozdaglar, 2009). Substantial evidence shows that individuals' beliefs are biased by location in a broad sense of the term: concrete geographical neighborhood or a more abstract positional descriptor such as culture, economic status, and social standing. For example, Das, Kuhnen, and Nagel (2017) show that a person's socio-economic status influences their beliefs about macroeconomic variables; Guiso, Sapienza, and Zingales (2006) describe the direct impact of culture on beliefs; Kuchler and Zafar (2018) find that when individuals form beliefs about aggregate house prices they overweight house-price observations from their local area; Shive (2010) and Bailey, Cao, Kuchler, and Stroebel (2018) find that investors' expectations are influenced by the experiences of other people within their social network.[2]

In our framework, each household regards its benchmark model as an approximation. Households believe that the data come from an unknown member of a set of models where firm-level expected returns differ from those in its benchmark model; i.e. each alternative model is characterized by an alternative probability measure. The household's concern

[1]Bipartite networks are a particular class of networks, whose nodes are divided into two sets, and only connections between nodes in different sets are allowed. For an introduction to the theory of networks, see the books by Easley and Kleinberg (2010) and Jackson (2010).

[2]There is also a large complementary literature showing that beliefs are influenced by personal experiences *over time* as opposed to location; see, for example, Vissing-Jorgensen (2003), Kaustia and Knüpfer (2008, 2012), Choi, Laibson, Madrian, and Metrick (2009), Greenwood and Nagel (2009), Chiang, Hirshleifer, Qian, and Sherman (2011), Malmendier and Nagel (2011, 2015), and Knüpfer, Rantapuska, and Sarvimäki (2017).

about model misspecification induces it to prefer decision rules that work over the set of alternative probability measures as opposed to one specific probability measure. At the same time, there is an "information penalty" for deviating from the benchmark probability measure. The household's information penalty depends on not just the alternative probability measure but also on its network location, in contrast with relative entropy.[3] For instance, if the household is located further from a particular firm, this reduces the size of the information penalty incurred for considering a given alternative probability measure, implying that the household is willing to consider larger deviations from the reference measure. Therefore, the further a household's location from a particular firm, the closer to zero will be the household's endogenous personal expectation for the firm's risk premium (i.e. the stock return in excess of the risk-free).

We use the network structure to represent the geographical location, which can be interpreted to include also cultural, social, and linguistic distance, of each household relative to firms and use this to derive the beliefs or probability measure of the household. Then, motivated by the "Universal Law of Generalization" developed in Shepard (1987), we model the effect of distance on beliefs via a negative exponential function. More recently, Sims (2018) has shown that "the universal law emerges inevitably from any information processing system (whether biological or artificial) that minimizes the cost of perceptual error subject to constraints on the ability to process or transmit information." Once we have obtained the endogenous beliefs of individual households, we derive the implications of these beliefs for the portfolio decisions of each household. Then, in our empirical analysis we exploit this theoretical link between beliefs, which cannot be observed, and portfolio holdings, which are observable, to test whether geography influences beliefs.

We consider a model with H heterogeneous households (each with its own beliefs) and N heterogeneous firms. As in Cox, Ingersoll, and Ross (1985), the physical capital of the N heterogeneous firms is subject to exogenous shocks. But, in contrast with Cox et al., we have heterogeneous households with Epstein and Zin (1989) and Weil (1990) preferences coupled with household-specific beliefs, as described above (and explained in detail below). We solve

[3]The relative entropy of an alternative probability measure with respect to a benchmark measure is the standard way of quantifying the change in information when an alternative probability measure replaces the benchmark measure. As such, relative entropy is based purely on information and ignores network location. If we were to ignore the influence of network location on the size of the information penalty, then our framework would reduce to that in Hansen and Sargent (2007).

this model in closed form and demonstrate how endogenous beliefs impact the portfolio choices of households. In our model, each household evaluates investment opportunities using its own personal beliefs, which depend on the household's location within the network of firms. Consequently, no household holds the market portfolio; instead, each household's portfolio is affected because of its personal beliefs. We then use this model along with data on the portfolio holdings of Finnish households to infer the effect of geographical distance on household beliefs about stock returns.

In particular, we test our model of belief formation using Finnish data on the portfolio holdings in 125 stocks of 405,628 households in 2,923 postal code areas. Using the portfolio holdings of Finnish households to infer their beliefs, we find that geographical distance between households and firms within the network has a statistically and economically significant effect on the beliefs of households about firms' stock returns. We find that the sensitivity coefficient representing distance is highly statistically significant in influencing beliefs in all regressions—univariate regressions, regressions including risk-aversion fixed effects, regressions including risk-aversion and stock-characteristic fixed effects, and panel regressions with robust standard errors double clustered at the firm and postal code level. The results are also *economically* significant: a one standard deviation decrease in distance to a firm's headquarters predicts an effect on beliefs that increases portfolio holdings by a factor of 2.645. We calculate the reduction in household welfare resulting from the deviation of beliefs from rational expectations and show how this varies depending on the location of households in the network.

A possible concern with these results is that they may be driven by households in Helsinki, the main conurbation in Finland, and there may be differences in behavior between rural and urban households. To check if this is indeed the case, we run regressions *excluding* stocks and postal codes in the Helsinki area and find that the results are still both statistically and economically highly significant. Another potential concern is that it may not be geographical distance per se that drives beliefs, but rather employment; i.e. households may tend to invest in the firms they work for, which are also likely to be close to where they live.[4] To investigate if this is the case, we *exclude* observations for which the household and firm headquarters are close to each other (under 8 miles in one specification

[4]See Cohen (2009) for empirical evidence on how loyalty to a company can influence portfolio choice.

and under 24 miles in another specification). We find that the results remain qualitatively similar, and thus, an "employment" effect does not seem to be driving the results.

Our paper is related to several streams of the literature. The first is the literature on robust decision making in finance and economics, which is described in Hansen and Sargent (2007). The key idea we take from this literature is that decision makers are uncertain about the benchmark model they use to make decisions. Consequently, they consider a range of models around the benchmark and make decisions that are robust with respect to the worst-case model. At the same time, there is a penalty for deviating from the benchmark model. This penalty is the relative entropy of the probability measure for each model that is considered with respect to the probability measure of the benchmark model. Trojani and Vanini (2004) study the implications of model uncertainty for portfolio choice in continuous-time economies with heterogenous investors. Gagliardini, Porchia, and Trojani (2008) provide an application of robust decision making in the Cox et al. (1985) model to study the implications for the term structure of interest rates. Bhandari, Borovička, and Ho (2019) develop a model where agents' subjective beliefs are endogenous consequences of model misspecification. Our approach extends this literature by using a penalty that incorporates network effects, in particular the location of a household relative to firms.

Our paper is therefore related to the literature on networks in economics and finance. A growing literature explores the implications of network structure for economic variables such as aggregate fluctuations (Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi, 2012, Gabaix, 2011), systemic risk and financial stability (Acemoglu, Ozdaglar, and Tahbaz-Salehi, 2015, Farboodi, 2017), systematic risk and asset pricing (Ahern, 2013), return predictability (Cohen and Frazzini, 2008), mortgage default risk (Stanton, Walden, and Wallace, 2014, 2018), merger waves (Ahern and Harford, 2014), information acquisition (Herskovic and Ramos, 2017), and asset pricing implications of information diffusion through networks (Walden, 2019). Allen and Babus (2009) and Jackson (2010, 2014) provide comprehensive surveys of this literature. Our work contributes to this literature by showing how networks can influence beliefs.

Finally, empirical tests of our model of belief formation rely on data on portfolio holdings of households. This aspect of our work is related to the seminal papers of Huberman (2001) and Grinblatt and Keloharju (2001), who document that households tend to overinvest in firms that are "familiar" to them. Huberman (2001) shows that familiarity depends on geographical distance, and Grinblatt and Keloharju (2001) find that it can depend also on cultural and language distance.[5] There is a large subsequent literature on "local bias" that documents the tendency of households to overinvest in companies "close" to them.[6] Our work contributes to this literature by providing a *belief-based microfoundation* for the local bias documented in these papers. We find empirical support for our model, by testing novel predictions that arise within our framework. For instance, the main regression tests in Grinblatt and Keloharju (2001) arise naturally in our framework, but their firm fixed effects now have particular economic interpretations. We show that the firm fixed effects in their tests represent distributional properties of stock returns (excess return per unit of variance risk), and postal-code fixed effects represent household preferences (risk aversion). Our belief-based model for their, and our, results is supported by the positive relation between the estimated excess return per unit of variance risk and out-of-sample firm performance. Novel empirical findings consistent with our model are the strong overlap between portfolio holdings of households within the same postal code and the strong relation between proximity and positive total portfolio holdings of a stock at the postal code level.

The rest of this paper is organized as follows. We describe the main features of our model in Section 2. The choice problem of an individual household with subjective beliefs is solved in Section 3. We evaluate the predictions of the model empiricallly in Section 4. We conclude in Section 5. Proofs for all results and additional empirical results are reported in the appendices.

[5] Lindblom, Mavruk, and Sjögren (2018) find that, in addition to local bias, individual investors who live in their birthplace invest almost three times more in local firms than other locals. Laudenbach, Malmendier, and Niessen-Ruenzi (2018) demonstrate that even decades after reunification, East Germans still invest significantly less in the stock market and are more likely to hold stocks of companies in communist countries (China, Russia, Vietnam), and less likely to invest in American companies and the financial sector.

[6] Massa and Simonov (2006) find evidence that households bias their portfolio toward stocks that are geographically or professionally close and argue that this bias is information driven. In contrast, Seasholes and Zhu (2010), Døskeland and Hvide (2011), Baltzer, Stolper, and Walter (2013, 2015) find that while portfolios are indeed biased toward local stocks, the local holdings do not generate abnormal performance. Bodnaruk (2009) provides compelling evidence of the effect of geographical distance by showing that as households change their place of residence, and thereby the distance from the companies in which they invest, they also adjust their portfolio weights. In contrast to the studies that focus on individual households, Coval and Moskovitz (1999a,b) study professional fund managers and find that they also bias their holdings toward local stocks and earn abnormal returns from nearby investments. Pool, Stoffman, and Yonker (2012) also study professional U.S. mutual-fund managers and find that they bias their portfolios toward stocks from their home states; however, home-state stocks do not outperform other holdings, suggesting that home-state investments are not informed.

2 The Model

In this section, we develop a model of a finite number of firms and households in a stochastic dynamic equilibrium economy. The location of a household relative to all firms in the economy is described by a network structure. The beliefs of a household are determined by its location within the network and the network itself.

2.1 Firms

There are N firms indexed by $n \in \{1, \ldots, N\}$. The value of the capital stock in each firm at date t is denoted by $K_{n,t}$ and the output flow by

$$Y_{n,t} = \alpha_n K_{n,t},$$

for some firm-specific technology level $\alpha_n > 0$. The level of a firm's capital stock can be increased by investment at the rate $I_{n,t}$. We thus have the following capital accumulation equation for an individual firm:

$$dK_{n,t} = I_{n,t}\, dt + \sigma_n\, K_{n,t}\, dZ_{n,t},$$

where σ_n, the volatility of the exogenous shock to a firm's capital stock, is constant. The term $dZ_{n,t}$ is the increment in a standard Brownian motion and is firm-specific; the correlation between $dZ_{n,t}$ and $dZ_{m,t}$ for $n \neq m$ is given by $0 < \rho < 1$, which is also assumed to be constant over time and the same for all pairs $n \neq m$. Firm-specific shocks create heterogeneity across firms. The $N \times N$ variance-covariance matrix of returns on firms' capital stocks is given by $V = [V_{nm}]$ and Ω denotes the correlation matrix, where the elements of these two matrices are

$$V_{nm} = \begin{cases} \sigma_n^2, & n = m, \\ \rho_{nm}\sigma_n\sigma_m, & n \neq m, \end{cases} \qquad \text{and} \qquad \Omega = \begin{cases} 1, & n = m, \\ \rho_{nm}, & n \neq m. \end{cases}$$

A firm's output flow is divided between its investment flow and dividend flow:

$$Y_{n,t} = I_{n,t} + D_{n,t}.$$

We can therefore rewrite the capital accumulation equation as

$$dK_{n,t} = \big(\alpha_n K_{n,t} - D_{n,t}\big)dt + \sigma_n\, K_{n,t}\, dZ_{n,t}.$$

In the Cox et al. (1985) model, the expected return on a firm's physical capital, α_n, equals the return on its stock. Similarly, the volatility of the return on a firm's capital, σ_n, equals the volatility of the return on its stock.

2.2 The Investment Opportunities of Households

There are H households, indexed by $h \in \{1, \ldots, H\}$. Households can invest their wealth in two classes of assets. The first is a risk-free asset, which has an interest rate i that we assume is constant over time, which then implies that the investment opportunity set is constant over time. Let $B_{h,t}$ denote the stock of wealth invested by household h in the risk-free asset at date t. Then, the change in $B_{h,t}$ is given by

$$\frac{dB_{h,t}}{B_{h,t}} = i\,dt.$$

Additionally, households can invest in the N risky firms, or equivalently, in the stocks of these N firms. We denote by $K_{hn,t}$ the stock of household h's wealth invested in the nth risky firm. Given that the household's wealth, $W_{h,t}$, is invested in the risk-free asset and the N risky firms, we have that:

$$W_{h,t} = B_{h,t} + \sum_{n=1}^{N} K_{hn,t}.$$

The proportion of a household's wealth invested in firm n is denoted by ω_{hn}, and so

$$K_{hn,t} = \omega_{hn} W_{h,t},$$

so that the amount of household h's wealth invested in the risk-free asset is

$$B_{h,t} = \Big(1 - \sum_{h=1}^{N} \omega_{hn}\Big) W_{h,t}.$$

The dividends distributed by firm n are consumed by household h according to share of firm n that household h holds:

$$C_{hn,t} = D_{hn,t} = \frac{K_{hn,t}}{K_{n,t}} D_{n,t},$$

where $C_{hn,t}$ is the consumption rate of household h from the dividend flow of firm n. Hence, the dynamic budget constraint for household h is given by

$$\frac{dW_{h,t}}{W_{h,t}} = \Big(1 - \sum_{n=1}^{N} \omega_{hn,t}\Big) idt + \sum_{n=1}^{N} \omega_{hn,t}\Big(\alpha_n dt + \sigma_n dZ_{n,t}\Big) - \frac{C_{h,t}}{W_{h,t}} dt,$$

where $C_{h,t}$ is the consumption rate of household h and $C_{h,t} = \sum_{n=1}^{N} C_{hn,t}$.

2.3 Network Structure

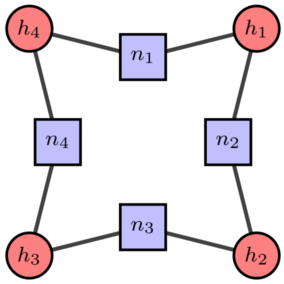
We endogenize household beliefs by linking them to the proximity of a household to each firm in the economy. That is, beliefs are the consequence of differences in locations across households, which we represent via a network structure. First, we explain the underlying network structure. Second, we show how household beliefs are modeled. Third, we show how household beliefs are impacted by their location within a network.

The separation of household h from firm n is denoted by d_{hn}. This could be a geographical distance or a more abstract measure of separation, such as cultural or linguistic distance. For example, an English-speaking household would be "more distant" from a German company located in Germany, than a German-speaking household that is located on the same street as the English-speaking household.

Shepard (1987) has proposed a "Universal Law of Generalization," which is one of the key psychological laws governing human cognition. This law states that the probability of distinguishing between two items, a and b, is a negative exponential function of the distance $d(a, b)$ between them in a psychological Euclidean space. This law is based on experiments where humans are presented with stimuli $S_a, S_b, \ldots$, about a set of items $a, b, \ldots$, and their response (typically having to identify the item) is given by $R_a, R_b, \ldots$. The stimulus S_a should evoke response R_a but with some probability can evoke response R_b where $b \neq a$; that is, item b is confused with item a. According to the universal law, the probability of this, $\Pr(R_b|S_a)$, is proportional to $e^{-d(a,b)}$. Shepard then shows that this law is consistent with empirical observations in a variety of contexts including having to distinguish among linguistic phonemes, circles of different sizes, and spectral hues. Motivated by these findings, we also specify a negative exponential function to model the effect of distance on portfolio choice.

Figure 1: First example of network

In this figure, we assume that the number of firms is equal to the number of households, $N = H = 4$, and that each household has a separation measure of $d \in (0, \overline{d})$ with respect to 2 firms and a separation measure of $\overline{d} + \epsilon$, where $\epsilon > 0$ with respect to all other firms.



We map the measure of separation, d_{hn}, into a measure of *proximity* using a negative exponential function, ϕ_{hn}, which is constrained to lie in the interval $[0, 1]$, by using the following specification

$$\phi_{hn} = \begin{cases} e^{-\kappa d_{hn}} & , d_{hn} \leq \overline{d}, \\ 0 & , d_{hn} > \overline{d}, \end{cases} \tag{1}$$

where $\kappa \geq 0$ is a measure of the sensitivity of ϕ_{hn} to d_{hn} and $\overline{d}$ is a constant denoting some threshold value. Thus, $\phi_{hn} = 1$ when the separation measure for household h relative to firm n is 0 and $\phi_{hn} = 0$ when the separation measure for household h relative to firm n exceeds the threshold $\overline{d}$. Note that both κ and $\overline{d}$ are common across all households.

In contrast with most existing work in finance, where a network consists of agents of one type, we have a network consisting of two types: households and firms. Such a network is known as a *bipartite* network. In our case, the bipartite network of households and firms is described via the H by N matrix $\mathbb{D}$, where

$$\mathbb{D} = [d_{hn}]_{hn}, \quad h \in \{1, \ldots, H\}, \, n \in \{1, \ldots, N\}.$$

The matrix $\mathbb{D}$ is called the biadjacency matrix of the bipartite network. From the biadjacency matrix of separation measures for household h, we obtain the proximity matrix

$$\Phi = [\phi_{hn}]_{hn}, \quad h \in \{1, \ldots, H\}, \, n \in \{1, \ldots, N\}.$$

Below, we provide two examples of bipartite networks. In our first example, illustrated in Figure 1, we assume the number of households equals the number of firms ($H = N$) and that each household has a separation measure of $d \in (0, \overline{d})$ with respect to 2 firms and a separation measure of $\overline{d}+\epsilon$, where $\epsilon > 0$ with respect to all other firms. Let the firms be arranged in a circle, and let each household h be equally distant from the two firms nearest to it on either side. Thus, in this case the biadjacency matrix of separation measures is given by the following N by N matrix

$$\mathbb{D} = \begin{bmatrix} d & d & \overline{d}+\epsilon & \overline{d}+\epsilon & \cdots & \overline{d}+\epsilon \\ \overline{d}+\epsilon & d & d & \overline{d}+\epsilon & \cdots & \overline{d}+\epsilon \\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \overline{d}+\epsilon & \overline{d}+\epsilon & \cdots & \overline{d}+\epsilon & d & d \\ d & \overline{d}+\epsilon & \cdots & \overline{d}+\epsilon & \overline{d}+\epsilon & d \end{bmatrix}.$$

Hence, defining $\phi = e^{-\kappa d}$, the matrix of proximity measures is given by

$$\Phi = \begin{bmatrix} \phi & \phi & 0 & 0 & \cdots & 0 \\ 0 & \phi & \phi & 0 & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \phi & \phi \\ \phi & 0 & \cdots & 0 & 0 & \phi \end{bmatrix}.$$

In our second example, observe Figure 2 below, which shows a map of Finland (in cyan) decomposed into 3036 postal code regions. (We will undertake a more detailed analysis of this data in our empirical tests in Section 4.) For the purpose of this example, consider the five red squares that represent five households situated at the points $p_1, \ldots, p_5$, and the three blue circles that represent three firms situated at p^1, p^2 and p^3.

Geographical distance is used to define the firm-household network, so that

$$d_{hn} = D((x_h, y_h), (x^n, y^n)) = \sqrt{(x_h - x^n)^2 + (y_h - y^n)^2},$$

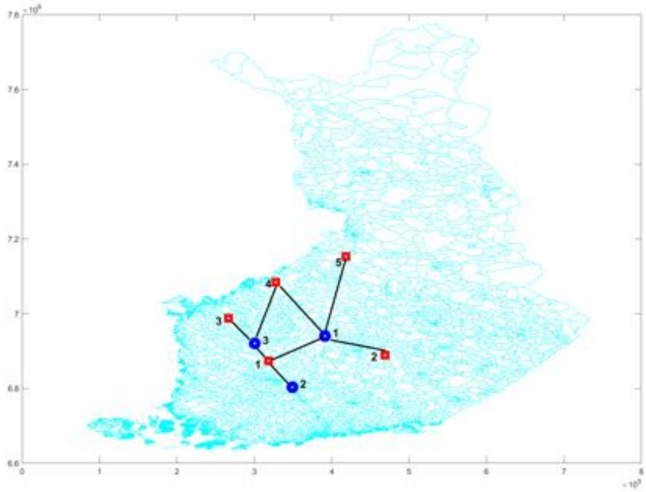
where the Euclidean distance function D is normalized so that the two postal codes that are farthest away from each other (postal codes 2 and 3) are at unit distance, and $p_h = (x_h, y_h)$, $p^n = (x^n, y^n)$.

The constant is $\kappa = 0.9$, and the threshold is set to $\bar{d} = 0.5$. This leads to the following matrix of proximity measures

$$\Phi = \begin{bmatrix} 0.771 & 0.852 & 0.905 \\ 0.770 & 0 & 0 \\ 0 & 0 & 0.850 \\ 0.720 & 0 & 0.736 \\ 0.676 & 0 & 0 \end{bmatrix}.$$

Figure 2: Second example of network

In this figure, we assume that the number of firms $N = 3$, with these firms represented by the three blue circles and the number of households is $H = 5$, with these households represented by the five red squares. The underlying map (in cyan), shows 3,036 Finnish postal code regions, obtained from the Finnish postal services company, Posti Group Corporation.



The proximity function is the highest for household 1 and firm 3 ($\phi_{13} = 0.905$), because they represent the geographically closest household-firm pair. There are in total eight household-firm connections; the remaining firm-household pairs are farther apart than the threshold $\bar{d} = 0.5$, and therefore, are not connected.

To show the flexibility of our framework, we also consider a variation of this example with a more general distance function. We now specify that the distance function is affected not only by geographical distance, but also cultural distance, as suggested in Grinblatt and Keloharju (2001). Culture could, for example, be measured by the main language spoken by the firm's CEO. Assume that the CEOs' main language is represented by $\ell^n \in \{0, 1\}$, where 0 represents a Swedish-speaking CEO and 1 a Finnish-speaking CEO. We shall assume that $\ell^1 = \ell^2 = 1$, and $\ell^3 = 0$. Also, assume similarly that $\ell_h \in \{0, 1\}$ represents the predominant language spoken by households in the different postal codes. We shall assume that $\ell_1 = \ell_4 = 0$, while $\ell_2 = \ell_3 = \ell_5 = 1$.

Then the generalized distance function takes the form $d_{hn} = d(p_h, p^n) + c|\ell_h - \ell^n|$, where the constant $c = 0.25$, leading to the following matrix of proximity measures

$$\Phi = \begin{bmatrix} 0 & 0.680 & 0.905 \\ 0.770 & 0 & 0 \\ 0 & 0 & 0.679 \\ 0 & 0 & 0.736 \\ 0.676 & 0 & 0 \end{bmatrix}.$$

With the generalized distance function, the connections between firm 1 and households 1 and 4 no longer exist, $\phi_{11} = \phi_{41} = 0$, because the effect of their cultural distance is to increase the distance beyond the threshold $\bar{d} = 0.5$. The connections ϕ_{12} and ϕ_{31} also decrease, from 0.852 to 0.680 and from 0.850 to 0.679, respectively, but stay positive, i.e. the connections remain.

2.4 Beliefs of Households

Each household h has its own beliefs, represented by its personal probability measure $\mathbb{Q}^{\boldsymbol{\nu}_h}$, which differs from the physical (objective) probability measure $\mathbb{P}$. We define the beliefs of household h below.

Definition 1. *If we consider an event A which can occur at time $T > t$, household h's personal expectation that event A could occur, conditional on date-t information, is given by*

$$E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}[I_A] = E_t^{\mathbb{P}}\left[\frac{M_{h,T}}{M_{h,t}} I_A\right], \tag{2}$$

where I_A is the indicator function associated with event A, and $M_{h,t}$ is an exponential martingale (the Radon-Nikodym derivative of $\mathbb{Q}^{\boldsymbol{\nu}_h}$ with respect to $\mathbb{P}$) defined by

$$\frac{dM_{h,t}}{M_{h,t}} = \boldsymbol{\nu}_{h,t}^{\top}(\Omega\Sigma)^{-1} d\boldsymbol{Z}_t,$$

where $\Sigma = diag(\sigma_1, \ldots, \sigma_N)$, $\boldsymbol{Z}_t = (Z_{1,t}, \ldots, Z_{N,t})^{\top}$, and

$$\boldsymbol{\nu}_{h,t} = (\nu_{h1,t}, \ldots, \nu_{hN,t})^{\top} \tag{3}$$

is the vector of divergences of household expectations from rational expectations.

To understand the intuition behind the above definition, observe that in equation (2) multiplication by the exponential martingale changes the expected rate of return for firm

n's stock from α under the measure $\mathbb{P}$ to $\alpha_n + \nu_{hn,t}$ under household h's measure, $\mathbb{Q}^{\boldsymbol{\nu}_h}$. Thus, the divergence in household h's beliefs about the expected rate of return on firm n (relative to the physical probability measure) is ν_{hn}. Consequently, household h's personal belief about the expected rate of return on its wealth is $\sum_{n=1}^{N} \omega_{hn,t}(\alpha_n + \nu_{hn,t})$ instead of $\sum_{n=1}^{N} \omega_{hn,t}\alpha_n$, a change of $\sum_{n=1}^{N} \omega_{hn,t}\nu_{hn,t}$ relative to the physical measure $\mathbb{P}$, which we can write more succinctly as $\boldsymbol{\omega}_{h,t}^{\top}\boldsymbol{\nu}_{h,t}$, where $\boldsymbol{\omega}_{h,t} = (\omega_{h1,t}, \ldots, \omega_{hN,t})^{\top}$ is the column vector of portfolio weights and $\boldsymbol{\nu}_{h,t}$, defined in equation (3), is the vector of personal divergences of household h for all N firms.[7] Only in the special case where household h's vector of personal divergences is the zero vector do its beliefs coincide with the objective beliefs represented by $\mathbb{P}$, and therefore, we have rational expectations.

In Definition 1, the only source of different beliefs between households is the vector of divergences, $\boldsymbol{\nu}_{h,t}$, which will be determined by a household's location. As a consequence, households in the same location will be predicted to hold identical portfolios. Moreover, as we shall see, households will neither overinvest nor shortsell stocks. These features are, of course, inconsistent with what is observed in practice, and are easily avoided with an extended specification. Specifically, we may replace $M_{h,t}$ in Definition 1 with the process

$$\frac{dM_{h,t}}{M_{h,t}} = (\boldsymbol{\nu}_{h,t} + \boldsymbol{\xi}_h)^{\top}(\Omega\Sigma)^{-1}d\boldsymbol{Z}_t,$$

where

$$\boldsymbol{\xi}_h = (\xi_{h1}, \ldots, \xi_{hN})^{\top},$$

and ξ_{hn} are random variables with mean zero, i.i.d. across agents and stocks. The additional variation in individual household's expectations introduced by this extension generates further heterogeneity in individual household beliefs and holdings, but cancels out in aggregate in markets with many households. For simplicity, we use the base specification in the forthcoming derivation; the derivation under the extension is very similar.

We now use the vector of household expectations to define the *network-weighted information loss* of an individual household.

[7] From Girsanov's Theorem, we know that choosing a vector of personal divergences is equivalent to a household changing the objective physical measure to a new measure, denoted by $\mathbb{Q}^{\nu_h}$.

Definition 2. *The network-weighted information loss for household h is given by*

$$\widehat{L}_{h,t} = \boldsymbol{\nu}_{h,t}^{\top}\Sigma^{-1}F_h\Sigma^{-1}\boldsymbol{\nu}_{h,t} = \sum_{n=1}^{N}\frac{\phi_{hn}}{1-\phi_{hn}}\frac{\nu_{hn,t}^2}{\sigma_n^2}, \tag{4}$$

where F_h is the $N \times N$ diagonal matrix

$$F_h = diag\left(\frac{\phi_{h1}}{1-\phi_{h1}},\ldots,\frac{\phi_{hN}}{1-\phi_{hN}}\right), \quad \phi_{hn} \in [0,1], \quad n \in \{1,\ldots,N\}. \tag{5}$$

To understand the motivation underlying the above definition, observe that in (4) we can interpret $\nu_{hn,t}^2/\sigma_n^2$ as a measure of the information about firm n that is discarded by household h when it uses its personal belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ as opposed to the objective belief $\mathbb{P}$. We then weight each of these information losses by $\frac{\phi_{hn}}{1-\phi_{hn}}$. Doing so ensures that an information loss impacts $\widehat{L}_{h,t}$ only when a household's proximity with respect to a particular firm is not zero and that the impact of the information loss increases with proximity and becomes infinitely large when $\phi_{hn} = 1$, i.e. when the separation measure is 0. In this manner, a household's location within a network structure determines its network-weighted information loss.

Our definition of a household's network-weighted information loss is closely related to the relative entropy per unit time of the objective belief $\mathbb{P}$ with respect to the personal belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ (also known as the Kullback-Leibler divergence from $\mathbb{Q}^{\boldsymbol{\nu}_h}$ to $\mathbb{P}$). Observe that the relative entropy per unit time of the objective belief $\mathbb{P}$ with respect to the personal belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ is given by

$$D^{KL}[\mathbb{P}|\mathbb{Q}^{\boldsymbol{\nu}_h}] = \frac{1}{dt}E_t\left[\left(\frac{dM_{h,t}}{M_{h,t}}\right)^2\right] = \boldsymbol{\nu}_{h,t}^{\top}\Sigma^{-1}[(\Omega)^{-1}]^{\top}\Sigma^{-1}\boldsymbol{\nu}_{h,t}. \tag{6}$$

In the above expression for relative entropy, we can see that when firm-level information losses are summed up, they are weighted by the correlation matrix for stock returns, Ω. In contrast, our definition of a household's network-weighted information loss in (4) uses the household's *location* in the bipartite network of firms and households, given by the matrix F_h defined in (5), to weight the information losses.

2.5 Intertemporal Aggregator with Endogenous Beliefs

Each household maximizes its date-t utility level, $U_{h,t}$, defined as in Epstein and Zin (1989) by an intertemporal aggregation of date-t consumption flow, $C_{h,t}$, and the date-t personal certainty-equivalent of date $t + dt$ utility:[8]

$$U_{h,t} = \mathcal{A}_h(C_{h,t}, \mu^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}]),$$

where $\mathcal{A}_h(\cdot,\cdot)$ is the time aggregator, defined by

$$\mathcal{A}_h(x, y) = \left[(1 - e^{-\delta_h dt})x^{1-\frac{1}{\psi_h}} + e^{-\delta_h dt}y^{1-\frac{1}{\psi_h}}\right]^{\frac{1}{1-\frac{1}{\psi_h}}}, \tag{7}$$

in which $\delta_h > 0$ is the rate of time preference, $\psi_h > 0$ is the elasticity of intertemporal substitution, and $\mu^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}]$ is the date-t personal certainty equivalent of $U_{h,t+dt}$.

We now use the concepts of a household's beliefs and its network-weighted information loss to define its personal certainty-equivalent, which is fundamental to how we endogenize belief formation as a function of household location within the network structure.

Definition 3. *The date-t personal certainty-equivalent of date-$t + dt$ household utility is given by*

$$\mu^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = \widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] + U_{h,t}L_{h,t}dt, \tag{8}$$

where $\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}]$ is defined by

$$u_{\gamma_h}\left(\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}]\right) = E^{\mathbb{Q}^{\boldsymbol{\nu}_h}}_t[u_{\gamma_h}(U_{h,t+dt})], \tag{9}$$

$$u_{\gamma_h}(x) = \frac{x^{1-\gamma_h}}{1-\gamma_h}, \quad \textit{and}$$

$$L_{h,t} = \frac{1}{2\gamma_h}\widehat{L}_{h,t}.$$

The first part of the definition of the personal certainty-equivalent in (8) is just the standard definition of a certainty-equivalent based on power utility with relative risk aversion γ_h, but using the personal belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$. The second part of the definition hinges on the expression for $L_{h,t}$, which is a multiple of the network-weighted information loss from using the personal belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ instead of the objective belief $\mathbb{P}$.

When a household chooses its beliefs by choosing a vector of divergences $\boldsymbol{\nu}_{h,t}$, it does so in order to minimize the impact of its information losses on its personalized certainty

[8] The only difference with Epstein and Zin (1989) is that we work in continuous time, whereas they work in discrete time. The continuous-time version of recursive preferences is known as stochastic differential utility (SDU), and is derived formally in Duffie and Epstein (1992). Schroder and Skiadas (1999) provide a proof of existence and uniqueness for the finite horizon case.

equivalent. The following proposition and corollary show how a household optimally selects its personal beliefs by choosing a vector of divergences, $\boldsymbol{\nu}_{h,t}$.

Proposition 1. *The date-t personal certainty equivalent of date-$t+dt$ household utility based on the personal belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ is given by*

$$\mu^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = \mu_{h,t}[U_{h,t+dt}] + U_{h,t} \times \left(\frac{W_{h,t} U_{W_{h,t}}}{U_{h,t}} \boldsymbol{\nu}^{\top}_{h,t} \boldsymbol{\omega}_{h,t} + L_{h,t} \right) dt, \tag{10}$$

where

$$\mu_{h,t}[U_{h,t+dt}] = E_t[U_{h,t+dt}] - \frac{1}{2}\gamma_h \, U_{h,t} \, E_t \left[\left(\frac{dU_{h,t}}{U_{h,t}} \right)^2 \right] \tag{11}$$

$U_{W_{h,t}} = \frac{\partial U_{h,t}}{\partial W_{h,t}}$ *is the partial derivative of household h's utility with respect to its wealth and E_t denotes the conditional expectation at t under the reference measure. At date t, a household optimally selects its personal belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ by choosing the vector of personal divergences $\boldsymbol{\nu}_{h,t}$ which minimizes its date-t personal certainty equivalent, $\mu^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}]$.*

The presence of the term $\frac{W_{h,t} U_{W_{h,t}}}{U_{h,t}} \boldsymbol{\nu}^{\top}_{h,t} \boldsymbol{\omega}_{h,t}$ in (10) gives a household the desire to make its divergences in portfolio-return expectations, i.e. $\boldsymbol{\nu}^{\top}_{h,t} \boldsymbol{\omega}_{h,t}$, as negative possible. This desire is clearly a departure from rationality, but it is tempered by the size of its network-weighted information losses, represented by $L_{h,t}$. For example, in the special case where $\phi_{hn} = 1$ for all $n \in \{1, \ldots, N\}$, the network-weighted information losses from any divergences become infinitely large, so a household chooses not to have any personal divergences, i.e. $\boldsymbol{\nu}_{h,t}$ is the zero vector. In this case, its personal certainty equivalent reduces to the standard certainty equivalent under rational expectations, i.e. a mean with a penalty for risk that depends on risk aversion γ_h and variance, as given in (11).

In general, network-weighted information losses are not infinitely large, and so a household faces a trade off between more negative personal divergences in portfolio return expectations and larger network-weighted information losses. Because this optimization problem is linear-quadratic it has the following closed-form solution.

Corollary 1. *A household's personal divergence vector is given by*

$$\boldsymbol{\nu}_{h,t} = -\gamma_h \frac{W_{h,t} U_{W_{h,t}}}{U_{h,t}} (\Sigma^{-1} F_h \Sigma^{-1})^{-1} \, \boldsymbol{\omega}_{h,t}, \tag{12}$$

with each element being

$$\nu_{hn,t} = -\gamma_h \frac{W_{h,t} U_{W_{h,t}}}{U_{h,t}} \sigma_n^2 \frac{1-\phi_{hn}}{\phi_{hn}}\, \omega_{hn,t}.$$

We can now see how network location affects households' expectations. When a household's proximity ϕ_{hn} to a firm is low, it responds by biasing downwards its point estimate for the expected return for that firm. For example, household h will change the expected return for firm n from α to $\alpha + \nu_{hn,t}$, thereby reducing the magnitude of the firm's expected risk premium ($\nu_{hn,t} \leq 0$ if $\omega_{hn,t} > 0$ and $\nu_{hn,t} \geq 0$ if $\omega_{hn,t} < 0$). The size of the reduction depends on each household's proximity to a particular firm—the reduction is smaller for firms to which the household is closer. Thus, differences in proximity ϕ_{hn} across households lead them to use different estimates of expected returns in their decision making.

3 Portfolio-Consumption Choice with Endogenous Beliefs

In this section, we solve the portfolio problem of an individual household whose beliefs are determined endogenously by its network location.

If a household's beliefs coincided with the objective physical measure, it would choose its consumption rate, $C_{h,t}$, and portfolio policy, $\boldsymbol{\omega}_{h,t}$, to solve the standard choice problem, which is:

$$\sup_{C_{h,t}} \mathcal{A}\Big(C_{h,t}, \sup_{\boldsymbol{\omega}_{h,t}} \mu_{h,t}[U_{h,t+dt}]\Big), \tag{13}$$

where $\mu_{h,t}[U_{h,t+dt}]$ is the standard certainty equivalent, given in (11).

In general, with endogenous household beliefs that do not coincide with rational expectations, the time aggregator $\mathcal{A}(\cdot)$ in (7) is unchanged—all we need to do is to replace the maximization of the standard certainty-equivalent in (13), $\sup_{\boldsymbol{\omega}_{h,t}} \mu_{h,t}[U_{h,t+dt}]$, with the combined maximization and minimization of the personal certainty equivalent,[9]

[9] Our max-min characterization of the objective function is consistent with the multi-prior approach advocated by Gilboa and Schmeidler (1989) and developed in a static setting by Dow and Werlang (1992), in dynamic discrete-time by Epstein and Wang (1994), and in continuous time by Chen and Epstein (2002). The alternative approach of Hansen and Sargent (2007), which our formulation builds on, assumes that investors allow for the possibility that their model may not be correct and hence consider deviations from the reference model, where the relative likelihood of the two models is measured using entropy; these preferences are called multiplier preferences. Maccheroni, Marinacci, and Rustichini (2006) show the relation between multiple-priors and multiplier preferences; they also show that both are nested in a larger class called "divergence preferences." Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011) provide a common representation that unifies these preferences.

$\sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} \mu^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}]$ to obtain

$$\sup_{C_{h,t}} \mathcal{A}\Big(C_{h,t}, \sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} \mu^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}]\Big). \tag{14}$$

A household, because of the impact of its network location, chooses $\boldsymbol{\nu}_{h,t}$ to *minimize* its personal certainty equivalent. By comparing (13) and (14), we can see that once a household has chosen the vector of personal divergences, $\boldsymbol{\nu}_{h,t}$, to adjust the expected returns of each firm to account for its beliefs, it makes consumption and portfolio choices in the standard way.

To solve a household's consumption-portfolio choice problem under subjective beliefs we use Ito's Lemma to derive the continuous-time limit of (14), which leads to a Hamilton-Jacobi-Bellman equation that is given in the appendix. The appendix also shows that the Hamilton-Jacobi-Bellman equation can be decomposed into a portfolio-optimization problem and an intertemporal consumption-choice problem. Given that the investment opportunity set is constant over time, the maximized household utility is a constant multiple of the household's wealth, which allows us to get the following simple expressions for the choice variables of the household.

Proposition 2. *The optimal vector of personal divergences is*

$$\boldsymbol{\nu}_h = -(I + VF_h\Sigma^{-2})^{-1}(\boldsymbol{\alpha} - i\mathbf{1}), \tag{15}$$

and the vector of optimal portfolio weights is

$$\boldsymbol{\omega}_{h,t} = \frac{1}{\gamma_h}V^{-1}\big(\boldsymbol{\alpha} - i\mathbf{1} + \boldsymbol{\nu}_h\big) = \frac{1}{\gamma_h}(V + \Sigma^2 F_h^{-1})^{-1}\big(\boldsymbol{\alpha} - i\mathbf{1}\big). \tag{16}$$

For the special case in which the correlation between assets $\rho_{nm} = 0$, the optimal divergence in household h's expected return for firm n is

$$\nu_{hn} = -(\alpha_n - i)(1 - \phi_{hn}), \tag{17}$$

and the optimal proportion of wealth invested in firm n by household h is

$$\omega_{hn} = \frac{1}{\gamma_h}\frac{\alpha_n - i + \nu_{hn}}{\sigma_n^2} = \frac{1}{\gamma_h}\frac{\alpha_n - i}{\sigma_n^2}\phi_{hn}. \tag{18}$$

From (17), we can see that the size of a household's deviation in a firm's expected return is smaller when the proximity measure, ϕ_{hn}, is larger; if $\phi_{hn} = 1$, then the deviation vanishes altogether and the household holds the portfolio weight that would be optimal under rational expectations. From (18), we see that the standard mean-variance portfolio weight for firm n, $\frac{1}{\gamma_h}\frac{\alpha_n - i}{\sigma_n^2}$, is scaled by the proximity measure for household h with respect to firm n, ϕ_{hn}. As a household's proximity measure with respect to a particular firm decreases, the proportion of its wealth that it chooses to invest in that firm also decreases.

In summary, when a household is further away from a firm (a reduction in proximity), its beliefs about the firm's return diverge more from the objective physical expectation; hence, it tilts its portfolio away from the one under rational expectations.

Our portfolio weight specification in equation (18) is similar to that estimated using panel regressions in Grinblatt and Keloharju (2001). Our specification provides additional insight about how to interpret the firm and location fixed effects in those regressions: The firm fixed effects represent distributional properties of stock returns, and the location fixed effects represent risk aversion, which we will measure at the postal code level. Our model thereby provides a microfoundation for the empirical tests and results in Grinblatt and Keloharju (2001).

It is straightforward to show that under the extended specification, which allows for idiosyncratic variation in household beliefs, equation (18) becomes

$$\omega_{hn} = \frac{1}{\gamma_h}\frac{\alpha_n + \xi_{nh} - i}{\sigma_n^2}\phi_{hn}.$$

This specification thus allows for variation in portfolio holdings across households at the same location (with the same ϕ_{hn}), with some choosing higher weights than $\frac{1}{\gamma_h}\frac{\alpha_n - i}{\sigma_n^2}$ (those with ξ_{hn} sufficiently high), and for some to shortsell stock (those with ξ_{hn} sufficiently low).

4 Empirical Results

Our belief-based portfolio-choice model leads to several testable predictions. These predictions follow immediately from our earlier results, and therefore, are presented below without formal proofs.

4.1 Testable Predictions

In our model, households will tend to be underdiversified in that they will hold only a subset of the stocks in the market (those at a distance smaller than $\bar{d}$). The effect will be especially severe for households that are located far away from firms.

Prediction 1. *The portfolio of a household located far away from firms is more underdiversified—i.e. contains fewer stocks—than the portfolio of a household located close to firms.*

A related prediction is that the stocks in a household's portfolio will be located closer to the household than stocks that are not part of the household's portfolio.

Prediction 2. *The distances from a household to firms that are included in the household's portfolio are lower than the distances to firms that are not included in the portfolio.*

The previous predictions relate a household's portfolio to the locations of firms. Our model also predicts that households which are located close to each other hold similar portfolios.

Prediction 3. *The portfolios of two nearby households contain more common stock constituents than if these constituents were randomly selected among the market's stocks.*

We next turn to the full portfolio implications of the specification in equations (1) and (18). Specifically, the sensitivity parameter κ, which is common across all households, determines how distance affects household beliefs. The higher is κ, the more households focus on companies in their network vicinity, whereas when $\kappa = 0$, as long as the distance is less than $\bar{d}$, beliefs and portfolios are unaffected.

We define $g_h = \ln(\gamma_h)$, $v_{hn} = \ln(\omega_{hn})$, and $s_n = \ln\left(\frac{\alpha_n - i}{\sigma_n^2}\right)$, which from (18) then implies

$$v_{hn} = -g_h + s_n - \kappa\, d_{hn}. \tag{19}$$

Equation (19) provides a strong characterization of the portfolio holdings of individual agents, as a function of relative risk aversion coefficients, stock return characteristics, and distance. Given data on portfolio holdings of the households in a market, $\{\omega_{hn}\}_{hn}$, together with location parameters, $\{(x_h, y_h)\}_h$ and $\{(x^n, y^n)\}_n$, equation (19) can be estimated and the following prediction can be tested

Prediction 4. *The coefficient* κ *in equation* (19) *is strictly positive,* $\kappa > 0$.

In this estimation, $\{g_h\}_h$ and $\{s_n\}_n$ are treated as (unobservable) household and firm fixed effects, respectively, both with clear economic interpretations. In particular, the estimated firm fixed effects should be related to the actual return distribution of stocks.

Prediction 5. *The firm fixed effects coefficients obtained when estimating equation* (19) *should be informative about*

$$\ln\left(\frac{\alpha_n - i}{\sigma_n^2}\right), \qquad n = 1, \ldots N.$$

Prediction 5 could potentially be used to separate our belief-based explanation for local bias from other reduced-form explanations, which would directly specify a positive relation between proximity and portfolio holdings. For example, a specification of portfolio holdings, $\omega_{h,n} = \mathsf{F}(\gamma_h, d_{h,n}, f_n)$, where γ_h represents individual household characteristics, and $d_{h,n}$ distance, and f_n firm characteristics unrelated to return distributions would not in general lead to the relation described in Prediction 5.

In the rest of this section we test Predictions 1–5. In Section 4.2, we describe the data. The results of our empirical analysis are reported in Section 4.3 and various robustness checks are described in Section 4.4.

4.2 Data

We obtained portfolio holdings for all accounts on the Helsinki Stock Exchange, as of January 2, 2003, from Euroclear, which acquired the Finnish Central Securities Depositary in 2008.[10] The data contain portfolio holdings and postal-code information, as well as further characteristics (age and gender and sector code classification) of all account holders in the market. There are altogether 3,036 valid postal codes in the data set and the data contains over 60 million trades during the time period 1995-2004, and about 1.2 million accounts, most of which represent the household sector.

We obtain geographical coordinates for each postal code area from the Finnish postal services company, Posti Group Corporation. These postal codes make up a fine-grained

[10] The dataset has previously been used, e.g., in Grinblatt and Keloharju (2000, 2001) and Walden (2019).

representation of Finland, as shown earlier in Figure 2. We represent each postal code geographically by its center of gravity.

We obtain information about the postal codes of company headquarter from Thomson One Reuters and exclude companies headquartered outside of Finland. We also exclude companies with shares that were not traded within the previous month (i.e. during December 2002). Finally, we exclude the telecommunications company Elisa Oyj, which had until 1999 been a privately held mutual association with broad ownership among its association members, and therefore had a quite different ownership history and structure than the rest of the firms.[11] This leaves us with 125 stocks, which are listed in Table D1 of Appendix D.[12]

We include accounts that are classified as households (these are accounts associated with sector codes between 500 and 599), that are associated with a valid postal code, and that owned shares in at least one of the 125 stocks on January 2, 2003. This leaves us with 405,868 households associated with altogether $P = 2,923$ postal codes.

The postal code associated with household h is denoted p_h. We assume that each agent resides at the center of gravity of his/her respective postal code and also that each firm is headquartered at the center of gravity of its postal code. Thus, all agents within a postal code are assumed to be at the same distance from each of the firms. We can then rewrite equation (19) at the postal-code level as

$$v_{pn} = -\bar{g}_p + s_n - \kappa\, d_{pn}, \tag{20}$$
$$\bar{g}_p = \ln(\bar{\gamma}_p), \qquad \frac{1}{\bar{\gamma}_p} = \sum_{\{h: p_h = p\}} \frac{1}{\gamma_h},$$

where $\bar{\gamma}_p$ is the harmonic mean of the relative risk aversion coefficients for agents living in postal code p, and d_{pn} is the distance between the center of gravity of postal code area p,

[11] Elisa Oyj was formed on July 1, 2000, with the merger of the Helsinki Telephone Corporation (in Finnish, "Helsingin Puhelin") and its holding company, HPY Holding Corporation. Helsinki Telephone Corporation had been a privately held telephone cooperative with broad ownership among its 550,000 "association members." Subscribers to its telephone services automatically became association members. When the company was listed on the Helsinki Stock Exchange in 1997, these owners became shareholders, which then carried over to Elisa Oyj after the merger (Source: Annual Reports, 1997-2000). Most of these shareholders only held this one stock, and thus seem to be shareholders for different reasons than the rest of the household investor population. Except for predictions 1 and 3, on underdiversification and similarity of holdings of individual households, the results are very similar when Elisa Oyj is included.

[12] Some of these stocks represent A and B shares in the same company. An A share in Finland typically come with greater voting rights compared with a B share. There were significant differences in share prices and returns between A and B shares of the same companies and we therefore include both A and B shares in our sample for companies with both types of shares.

(x_p, y_p), and the center of gravity of the postal code area in which firm n is located,

$$d_{pn} = D((x_p, y_p), (x^n, y^n)).$$

Equation (20) thus consists of $P \times N$ postal code/company holdings.

We focus on geographical proximity as a measure of network proximity between agents and companies. As discussed in Walden (2019), this is a reasonable assumption for the time period and market that the data covers. One might view geographical distance as an unimportant hurdle in the present time due to almost universal access for households to information via the Internet. However, only about one third of the Finnish population used the Internet in 2000. Moreover, over half of Finland's population resides in rural areas, making it one of the most rural countries in the European Union. It is therefore plausible that there would be a significant link between geographical and network proximity in the early 2000's.

We normalize the distance function, so that all geographical coordinates lie in the unit square, $[0, 1] \times [0, 1]$. The household-firm that are farthest apart are therefore at a distance somewhere between 1 and $\sqrt{2}$ from each other (in our sample, the maximum distance is 1.175), making the interpretation of the sensitivity coefficient κ in equation (20) straightforward.

The geographical coordinates of postal codes (blue small dots) and firms (red large circles) are shown in Figure 3. As can be seen in the figure, most firms are headquartered in the far south (around the capital, Helsinki, with associated postal codes between 100 and 9900), whereas about 20% of the firms have headquarters elsewhere. Summary statistics of the data we use are provided in Table 1.

4.3 Results

We first test Prediction 1, that households tend to be more under-diversified the farther away they are located from stocks. We define the center of gravity (CoG) of the stocks,

$$(x^C, y^C) = \frac{1}{\sum_{n=1}^N W^n} \left(\sum_{n=1}^N W^n x^n, \sum_{n=1}^N W^n y^n \right),$$

where we use both equally-weighted CoG ($W^n = 1$ for all firms) and value-weighted CoG, (a firm's weight is defined as the total value of household portfolio holdings in that firm).

Figure 3: Center of gravity of postal codes for households and firms

Center of gravity of postal codes for households (blue dots) and firms (red circles) in dataset.

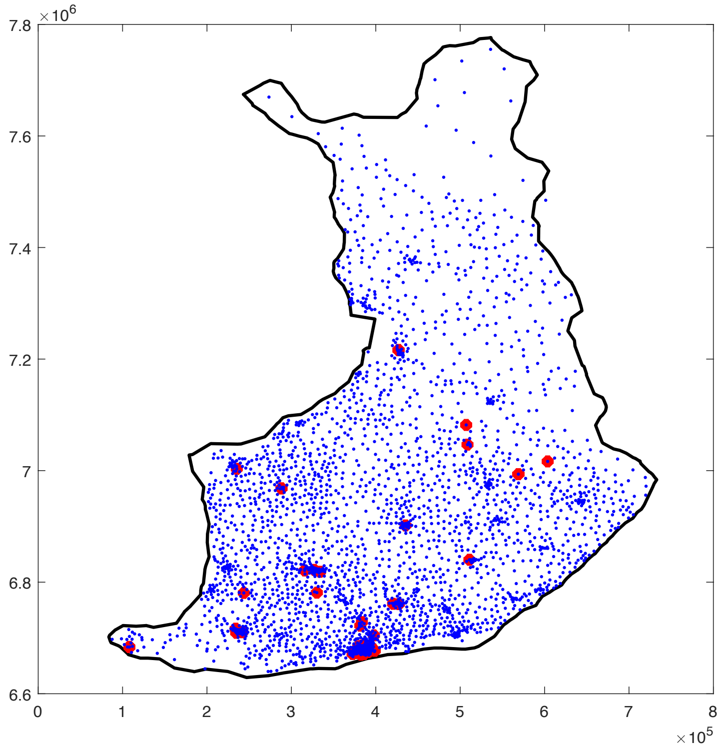


Table 1: Summary statistics

This table gives the summary statistics for the data used in our empirical analysis.

Number of stocks	125
Number of household accounts	405,628
Number of postal codes	2,923
Average number of accounts in postal code	139
Total number of observations	364,750
Number of nonzero observations	132,811
Maximum portfolio holding	FIM 178.7 million
Minimum portfolio holding	FIM 0
Average stock holding by postal code	FIM 30,617
Median number of stocks held by account	2
Mean number of stocks held by account	2.75

We further define q_p, the average number of stocks in households' portfolios within postal code area p, and $d_p = D((x_p, y_p), (x^C, y^C))$, the distance between households in postal code area p and firm CoG, and estimate how they are related. The relation, shown in Table 2,

Table 2: Test of Prediction 1: Underdiversification versus distance from firms

This table shows how the average number of stocks held by households in a postal code are is related to the distance of the postal code to firm center-of-gravity (CoG), d_p. Columns (1) and (2) use an equally-weighted definition of CoG, whereas columns (3) and (4) use a value-weighted definition. Univariate regressions are used in columns (1) and (3), and bivariate regressions, also including log-average portfolio size, $\ln(W_p)$, are used in columns (2) and (4).

	(1)	(2)	(3)	(4)
Distance, d_p	−1.011***	−0.306***	−1.024***	−0.313***
Standard error	0.084	0.072	0.082	0.070
Portfolio size, $\ln(W_p)$		0.446***		0.445***
Standard error		0.012		0.012

is significantly negative, both economically and statistically. A household located at a maximum distance from the firms' CoG is predicted to hold about one fewer stock in its portfolio than a household located right at the CoG—a major effect because the median number of stocks held in a household portfolio is only 2. The results are very similar regardless of whether the value-weighted or equally weighted definition of CoG is used.

We also control for portfolio size. Specifically, if wealthier households tend to be better diversified and are also located closer to firms on average, similar results would arise, but the results would be because of this omitted variable. We therefore include the logarithm of average portfolio-size within a postal code area, $\ln(W_p)$, in the regression. The estimated coefficient decreases by about two thirds when $\ln(W_p)$ is included, both in the value-weighted and equally weighted CoG specification, but still remains highly significant.

We next test Prediction 2, which at the postal code level states that postal codes with zero portfolio holdings in a firm tend to lie geographically further away from that firm than postal codes with positive holdings. For each stock, we perform a two-sample t-test, comparing the average distances of postal codes with zero and with positive holdings from the firm's headquarter. The results are shown in Table 3.

The average distance for postal codes with zero holdings is about 0.332, whereas the average distance for postal codes with positive holdings is about 0.224, corresponding to a difference of 0.108—about 85 miles. At the individual firm level, the average distance from postal code with zero holdings is higher for 119 of the 125 stocks. For 116 of these firms, the difference is statistically significant at the 0.01% level. For the remaining 6 firms,

Table 3: Test of Prediction 2: Zero versus positive portfolio holdings

This table shows results from two-sample t-tests of difference in means between the distance from postal codes with zero holdings to firm headquarter and from postal codes with positive holdings.

	Zero holding	Positive holding	Difference
Average distance to HQ	0.332	0.224	0.108
Number of firms	6	119	
Average t-stat	−3.047	14.080	
Average t-stat, Total			13.258

for which the average distance is higher from postal codes with positive holdings, only one is significant at the 0.01% level. The average t-statistic for the difference of means being positive is 13.3. Thus, the data strongly support the prediction that the further away from a firm's headquarters, the higher the likelihood of zero portfolio holdings.

We also estimate the cutoff distance, $\bar{d}$. Specifically, for each of the 125 stocks, we choose the $\bar{d}$ that maximizes the number of correctly classified postal codes with respect to whether the stock holdings in the firm is positive or zero. The results are shown in Table 4. The average estimated $\bar{d}$ is 0.1799, corresponding to a threshold distance of about 143 miles. Beyond this distance to a firm's headquarter, the familiarity is thus predicted to be so low that an investor completely avoids investing in a stock.

The fraction of holdings that are zero among all postal code/firm observations is slightly less than two thirds, about 64%. When the estimated cutoff thresholds at the firm level are used to predict whether a postal code/firm portfolio holding is zero, the fraction of correct classifications is about 75%, i.e. about three quarters. The model thus captures quite well whether investors in a postal code invest in a stock or not.

Prediction 3, formulated at the postal code level, suggests that that households located in the same postal code should hold portfolios with more overlap (i.e. more common stocks) than if the portfolios were randomly chosen. For example, for two households that each randomly and independently invest in one of N stocks, with probability $\frac{1}{N}$ of choosing each stock, the probability that the two portfolios overlap would also be $\frac{1}{N}$. In other words, the expected value of the random value $\tilde{N}^c$, which denotes the number common stocks in their portfolios, would be $\frac{1}{N}$. More generally, if household h_1 randomly chooses N_{h_1} stocks and

Table 4: Estimated threshold for zero portfolio holdings

This table shows results the estimated thresholds, $\bar{d}$, for when portfolio holdings become zero.

	Value
Average threshold, $\bar{d}$	0.1799
Standard deviation	0.2165
Fraction of zero holdings	0.6359
Fraction correctly classified	0.7534

household h_2 randomly and independently chooses N_{h_2} stocks, then the expected overlap is

$$E\left[\tilde{N}^c_{h_1,h_2}\right] = \frac{N_{h_1}N_{h_2}}{N}.$$

Prediction 3 can then be tested by studying the actual versus expected overlap of the portfolios of all households located within the same postal code area, and sum up these overlaps over all postal codes, i.e. by calculating the actual total overlap

$$N^c_{TOT} = \sum_p \sum_{\substack{h_1,h_2\in\mathcal{H}_p \\ h_1\neq h_2}} N^c_{h_1,h_2},$$

where $\mathcal{H}_p = \{h : p_h = p\}$ is the set of households in postal code area p, and comparing it with expected total overlap.

The actual total overlap in the data is $N^c_{TOT} = 51,604,247$, whereas the expected total overlap under the independent, $\frac{1}{N}$ probability-per-stock, assumption is $3,478,029$. Thus, the likelihood that two randomly chosen households within the same randomly chosen postal code area invest in the same stock is almost 15 times higher than expected.

A limitation of the above test is that it assumes that each stock is chosen with the same probability, $\frac{1}{N}$, whereas in practice some stocks (for example, the telecommunications company Nokia in our dataset) are much more broadly held than others. The higher overlap will therefore partly be a consequence of the $\frac{1}{N}$ assumption. We therefore use a more sophisticated bootstrapping method, based on the actual distribution of portfolio holdings among the households to test the prediction. Briefly, we assume that the empirical distribution of household portfolio holdings represent the underlying data generating process, and compare the expected overlap if households and their portfolios were randomly assigned to postal codes with what is observed in the data. An interesting property of this test is that the maximum portfolio size, measured by number of stocks in the portfolio, can be capped, so that

Table 5: Test of Prediction 3: Overlap of portfolios within a postal code area

This table tests whether portfolio overlaps within a postal code area is higher than expected if households were randomly assigned to the area. Column (1) shows the maximum portfolio size included in test, column (2) the number of households included in the sample, column (3) the expected number of overlaps, column (4) the actual number of overlaps, and column (5) the ratio between actual and expected overlaps.

Max. portf. size	Number of households	Actual overlap	Expected overlap	Ratio
1	188,535	6,324,570	2,649,371	2.387
2	268,268	12,989,232	6,096,314	2.130
5	357,772	28,450,235	16,782,239	1.695
10	393,606	42,315,889	27,935,744	1.515
20	404,277	49,978,463	34,481,911	1.449
125	405,628	51,604,247	35,952,390	1.436

the test is applied to a subpopulation of households whose portfolios contain no more than 1, 2, 5, etc., stocks. Because location, via personal beliefs, drives both underdiversification and the choice of similar portfolios for nearby households, we expect higher-than-expected overlaps to be more pronounced among households that hold few stocks.[13] Further details of the test are provided in Appendix D.

The results of the bootstrapping method are shown in Table 5. The ratio of actual-to-expected overlap (the right-most column in the table) is the highest, 2.387, when only households holding one stock are included in the test, and then gradually decreases as the maximum portfolio size increases. The ratio is 1.436 when the full sample of households is included. These ratios are thus lower than the ratio based on the $\frac{1}{N}$ assumption, but remain highly significant, in support of Prediction 3.

Next, we test Prediction 4, our empirical specification of belief formation in equation (20), which can be viewed as a panel with fixed effects for postal codes (representing the log-risk aversion coefficients, g_p) and companies (representing the log-return characteristics, s_n). We therefore use panel regressions.

The model is not completely identified because the mapping $g_h \mapsto g_h + c$, $s_n \mapsto s_n + c$ yields the same household portfolio weights for an arbitrary constant c. Intuitively, higher risk aversion coefficients are offset by more favorable investment opportunities. We obtain

[13]There is also a purely mechanical effect in that the actual overlap cannot be much higher than the expected overlap when households hold large portfolios. For example, if a household holds *all* stocks in its portfolio, *every* stock in the other household's portfolio is common for the two portfolios, so actual and expected overlap must coincide in that case.

Table 6: Test of Prediction 4: Estimated sensitivity coefficient, κ

This table estimates the sensitivity coefficient, κ. Panel A includes observations with $\omega_{pn} = 0$ replaced with $\omega_{pn} = 1$ (one Finnish Mark, corresponding to about USD 0.17). Panel B excludes observations with $\omega_{pn} = 0$. Univariate in column 2, including risk aversion in column 3, including risk aversion and stock distributions in column 4, panel regression with postal code and stock fixed effects and robust standard errors double-clustered at the postal code and firm level in column 5. Statistical significance levels: * = 0.01, ** = 0.001, *** = 0.0001.

(1)	(2)	(3)	(4)	(5)
Panel A				
Sensitivity coefficient, κ	5.873***	3.861***	3.188***	3.185***
Standard error	0.037	0.058	0.056	0.434
log risk aversion, g				
-average		-4.330	-4.120	
-max		-0.382	-0.325	
-min		-12.887	-12.834	
log distribution, s				
-average			0	
-max			7.258	
-min			-2.658	
R^2	0.065	0.385	0.591	0.591
Adj. R^2	0.065	0.380	0.588	0.588
$N = 368,298$				
Panel B				
Sensitivity coefficient, κ	2.251***	2.123***	2.669***	2.669***
Standard error	0.035	0.061	0.046	0.403
log risk aversion, g				
-average		-8.572	-7.979	
-max		-4.563	-2.638	
-min		-12.845	-12.883	
log distribution, s				
-average			0	
-max			4.166	
-min			-3.985	
R^2	0.030	0.210	0.626	0.626
Adj. R^2	0.030	0.191	0.618	0.618
$N = 134,902$				

unique identification by normalizing the results so that the average s_n coefficient is 0 ($s_n = 0$ is, for example, obtained with $\alpha_n - i = 0.09$, $\sigma_n = 0.3$).

Finally, we consider two approaches for handling portfolio weights of zero, for which the logarithm is not defined. There are a large number of such observations in the data, even with aggregation at the postal code level. Under the first approach, we include these observations but replace the zero with a small positive threshold, namely one Finnish Mark (corresponding to about USD 0.17). Under the second approach, we exclude such observations and run unbalanced panel regressions. The disadvantage of the former approach is that it introduces an arbitrary lower threshold, whereas the disadvantage of the latter is that it does not use information about zero holdings.

The results are shown in Table 6. We see that the coefficient measuring sensitivity to distance, κ, is highly significant in all regressions (univariate, including risk-aversion fixed effects, including risk-aversion and stock-characteristic fixed effects, and panel regression with robust standard errors double clustered at the firm and postal-code level).

The results are also economically significant. The standard deviation of the distance between headquarter and household is 0.312. For the coefficient estimate that includes risk-aversion, stock distribution, and all observations (first row in Panel A, column 4, $\kappa = 3.118$), a one standard deviation decrease in distance to a firm's headquarters predicts an increase in portfolio holdings by a factor $e^{3.1880 \times 0.312} = 2.645$. The R^2 for the univariate regression in Panel A is 0.0654, corresponding to a correlation between network proximity and log-portfolio holdings of about 0.26.

Finally, we test Prediction 5, whether the estimated s_n coefficients (the firm fixed effects) from the belief-based model are informative about stocks return distributions out-of-sample. We calculate daily mean excess returns, $\hat{\alpha}_n - i$, and volatility, $\hat{\sigma}_n$, over a five-year period, from 2003-2005.

The results are shown in Table 7. We first compare the relation between estimated s_n and realized volatility, $\ln(\hat{\sigma}_n)$ because realized volatility is the part of $\ln\left(\frac{\hat{\alpha}_n - i}{\hat{\sigma}_n^2}\right)$ that is easiest to estimate. This relation should be negative. As shown in Column (1) in Panel A of the table, the relation is strongly negatively significant, with an R^2 of over 50%. We compare this with the prediction of the rational-expectations model. In the rational-expectations model, all households choose the same risky portfolio, $\phi_{hn} \equiv 1$ for all h and n, and by summing (19) over households, it follows that s_n is directly related to a firm's log-size. Column (2) in Panel A of Table 7 shows that firm log-size also is informative about volatility in our data, but with lower explanatory power than the belief-based model. Moreover, when both the belief-based and rational-expectations estimates are included in a bivariate

Table 7: Test of Prediction 5: Predictive power of estimated s_n coefficients

This table tests whether the s_n coefficients estimated from portfolio holdings predict realized volatility (Panel A) and return over variance (Panel B) in subsequent 5-year period. Column (1) uses the belief-based model, column (2) the rational expectations model, and column (3) both models. Realized volatility and returns measured using daily data over 3-year period, 2003-2005. Statistical significance levels: * = 0.01, ** = 0.001, *** = 0.0001.

	(1)	(2)	(3)
Panel A: $\ln(\hat{\sigma}_n)$			
Rational expectations, s_n^{RE}		−0.149***	0.036
Belief-based, s_n^{BB}	−0.185***		−0.219***
R^2	0.506	0.356	0.510
N	125		
Panel B: $\ln\left(\frac{\hat{\alpha}_n - i}{\hat{\sigma}_n^2}\right)$			
Rational expectations, s_n^{RE}		0.075	−0.296*
Belief-based, s_n^{BB}	0.149**		0.426***
R^2	0.068	0.019	0.121
N	108		

regression, as done in Column (3) of Panel A, the belief-based estimate dominates, remaining significant, whereas the log-firm size coefficient switches sign and becomes insignificant.

In Panel B of the table, we do the same estimation for $\ln\left(\frac{\hat{\alpha}_n - i}{\hat{\sigma}_n^2}\right)$ as we did for $\ln(\hat{\sigma})$, including the 108 stocks that had positive excess realized returns during the period (for which the logarithm is defined). The explanatory power is much lower both for the belief-based coefficient and the coefficient based on the rational expectation model, due to the well-known challenges of estimating stocks' expected returns from realized returns. However, the belief-based estimate remains significant at the 0.1% level in the univariate regression, and at the 0.01% level in the bivariate regression, while the rational-expectations coefficient switches sign in the bivariate regression. As a robustness check, we also run the same tests using three years of realized returns. The results (not reported) are similar, the main difference being that the belief-based coefficient in the bivariate regression in Panel B is significant only at the 10% level.

We stress that only data on portfolio holdings and location is used in the estimation of the s_n's. That these are informative about out-of-sample stock-return distributions therefore lends significant support for our belief-based explanation of observed local bias.

4.4 Robustness

A possible concern with the previous results is that they may be driven by different behaviors of households in urban areas (specifically, around Helsinki) relative to rural areas. For example, households in the Helsinki area may have a preference for stocks headquartered in an urban area for some other reason than network proximity, which will then lead to results similar to those reported above. To check if this is indeed the case, we run the regressions *excluding* stocks and postal codes in the Helsinki area (postal codes with fewer than 5 digits). The results are reported in Table D2 in Appendix D. We see that the results are qualitatively similar as before, and specifically, still both statistically and empirically highly significant. We also exclude households in the Helsinki area but not stocks, with similar results (not reported).

A potential alternative explanation for the results is that it may not be geographical distance per se that drives beliefs, but rather employment. That is, if households tend to invest in the firms they work for—which they likely also live close to—similar results may arise. To rule out such an explanation, we exclude observations for which the postal codes of account holder and firm headquarter are close. Specifically, we exclude all observations for which the normalized distance is less than some d_0. Table D3 in Appendix D, shows that the results remain qualitatively similar when $d_0 = 0.01$ (corresponding to a minimal distance of about 8 miles between postal code of account holder and firm headquarter for an observation to be included), and $d = 0.03$ (corresponding to a minimal distance of about 24 miles). Thus, an employment effect does not seem to be driving the results.

Our approach provides a tightly specified model that links "local bias" to familiarity, through adjustments of investors' beliefs about a firm as a function of distance—in our empirical specification captured by geographical distance. Other explanations for local bias that have been put forward in the literature are transaction costs and hedging demand.

We argue that transaction costs are unlikely to play a major role in explaining home bias within a country, especially when all stocks are traded on the same exchange. With respect to hedging demand, an alternative explanation for why agents prefer to invest in nearby stocks is that they provide a hedge against local shocks. For example, if a local firm performs well, prices of services and goods (e.g., housing) may increase because of increased

Table 8: Estimate of sensitivity coefficient with additional controls

This table estimates the sensitivity coefficient, κ, when including investor year of birth (YOB) and gender (GEN). Panel A sets $d_0 = 0$, whereas Panel B sets $d_0 = 0.05$, corresponding to a threshold distance of about 40 miles. No fixed effects in column 2, including risk aversion in column 3, including risk aversion and stock distributions in column 4, panel regression with postal code and stock fixed effects and robust standard errors double-clustered at the postal code and firm level in column 5. Statistical significance levels: * = 0.01, ** = 0.001, *** = 0.0001.

(1)	(2)	(3)	(4)	(5)
Panel A				
Sensitivity coefficient, κ	2.298***	2.056***	2.658***	2.658***
Standard error	0.034	0.060	0.046	0.398
Year of birth coefficient, YOB	−0.031***	−0.040***	−0.013***	−0.013***
Standard error	0.001	0.001	0.001	0.001
Gender coefficient, GEN	0.597***	0.445***	−0.157***	−0.157***
Standard error	0.019	0.018	0.013	0.033
R^2	0.065	0.249	0.630	0.630
Adj. R^2	0.065	0.233	0.622	0.622
$N = 134,825$				
Panel B				
Sensitivity coefficient, κ	0.809***	0.704***	1.377***	1.377***
Standard error	0.040	0.073	0.056	0.263
Year of birth coefficient, YOB	−0.031***	−0.037***	−0.012***	−0.012***
Standard error	0.001	0.001	0.001	0.001
Gender coefficient, GEN	0.461***	0.345***	−0.169***	−0.169***
Standard error	0.019	0.019	0.014	0.033
R^2	0.039	0.200	0.606	0.606
Adj. R^2	0.039	0.180	0.596	0.596
$N = 115,256$				

demand from the employees at the firm who are now wealthier. Investing in the local firm provides a hedge against such price shocks. Inasmuch as such hedging demand is related to the age and gender of the population, we can assess its affect, because account level information is available in the data. For example, younger investors—who are less likely to own their home—are likely to be more exposed to real estate price shocks than older investors.

We create variables for the average birth year (YOB) of the investors in a specific stock and postal code, and for their gender (GEN, which is 1 for male and 2 for female). The correlations between the distance to firm headquarter and these variables are both low,

$\rho_{d,\text{YOB}} = -0.049$, and $\rho_{d,\text{GEN}} = -0.056$, suggesting that any hedging demand that varies with age and/or gender is not captured by the distance to firm headquarter. To further explore a potential relation, we rerun the tests from Table 6, Panel B, but including birth year and gender, and also from Table D3 with $d_0 = 0.05$, to rule out hedging demand for local shocks within a 40-mile radius.

As seen in Table 8, the coefficient estimates for κ barely change and are still highly significant. We conclude that it is unlikely that our results are driven by hedging demand against local shocks, at least within a 40-mile radius of the investor, and against shocks that are related to age and/or gender.

A concern may be that headquarter provides a very rough measure of a firm's location. For example, some firms have operations spread out over the whole country and will be familiar to households far away from its headquarter. The challenges of developing an objectively superior alternative measure of firm location are significant though, which is why we use the well-established headquarter measure of location. As a robustness check, we ensure that the results do not change when excluding the two "least local" companies during the period: the global telecommunication company Nokia, which made up over half of the stock market value in the early 2000s, and the retail store chain Stockmann, which had stores all over Finland. The results (not reported) are very similar when excluding those two companies.

5 Conclusion

Motivated by empirical evidence that rejects the rational expectations hypothesis, we develop a model where a household's beliefs are an *endogenous* outcome of its location in a bipartite network of households and firms. We then evaluate the model empirically using data on portfolio holdings to infer household beliefs. The empirical evidence indicates that geographical distance between the locations of households and firms has a statistically and economically significant effect on the beliefs of households about stock returns. We calculate the reduction in household welfare resulting from the deviation of beliefs from rational expectations and show how this varies depending on the location of households in the network.

A Proofs

In this appendix, we provide all derivations for the results in the main text.

Proof of Proposition 1

The definition of the certainty equivalent in (9) implies that

$$\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[U_{h,t+dt}^{1-\gamma_h}\right]^{\frac{1}{1-\gamma_h}}.$$

Therefore

$$\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[U_{h,t+dt}^{1-\gamma_h}\right]^{\frac{1}{1-\gamma_h}} = E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[U_{h,t}^{1-\gamma_h} + d(U_{h,t}^{1-\gamma_h})\right]^{\frac{1}{1-\gamma_h}}.$$

Applying Ito's Lemma, we obtain

$$\begin{aligned} d(U_{h,t}^{1-\gamma_h}) &= (1-\gamma_h)U_{h,t}^{-\gamma_h}dU_{h,t} - \frac{1}{2}(1-\gamma_h)\gamma_h U_{h,t}^{-\gamma_h-1}(dU_{h,t})^2 \\ &= (1-\gamma_h)U_{h,t}^{1-\gamma_h}\left[\frac{dU_{h,t}}{U_{h,t}} - \frac{1}{2}\gamma_h\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right]. \end{aligned}$$

Therefore

$$\begin{aligned} \widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[U_{h,t+dt}^{1-\gamma_h}\right]^{\frac{1}{1-\gamma_h}} &= U_{h,t}\left(E_t\left[1+(1-\gamma_h)\left[\frac{dU_{h,t}}{U_{h,t}} - \frac{1}{2}\gamma_h\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right]\right]\right)^{\frac{1}{1-\gamma_h}} \\ &= U_{h,t}\left(1+(1-\gamma_h)\left[E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[\frac{dU_{h,t}}{U_{h,t}}\right] - \frac{1}{2}\gamma_h E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right]\right]\right)^{\frac{1}{1-\gamma_h}} \\ &= U_{h,t}\left(1+(1-\gamma_h)\left[E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[\frac{dU_{h,t}}{U_{h,t}}\right] - \frac{1}{2}\gamma_h E_t\left[\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right]\right]\right)^{\frac{1}{1-\gamma_h}}. \end{aligned}$$

Hence,

$$\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = U_{h,t}\left(1+E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[\frac{dU_{h,t}}{U_{h,t}}\right] - \frac{1}{2}\gamma_h E_t\left[\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right]\right) + o(dt).$$

Therefore, in the continuous time limit, we obtain

$$\frac{\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[dU_{h,t+dt}]}{dt} = \frac{\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] - U_{h,t}}{dt} = U_{h,t}\left(\frac{1}{dt}E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[\frac{dU_{h,t}}{U_{h,t}}\right] - \frac{1}{2}\gamma_h\frac{1}{dt}E_t\left[\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right]\right).$$

From Girsanov's Theorem

$$E_t^{\mathbb{Q}^{\boldsymbol{\nu}_h}}\left[\frac{dU_{h,t}}{U_{h,t}}\right] = E_t\left[\frac{dU_{h,t}}{U_{h,t}}\right] + E_t\left[\frac{dU_{h,t}}{U_{h,t}}\frac{dM_{h,t}}{M_{h,t}}\right].$$

Therefore

$$\begin{aligned}\frac{\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[dU_{h,t}]}{dt} &= \frac{\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] - U_{h,t}}{dt} \\ &= U_{h,t}\left(\frac{1}{dt}E_t\left[\frac{dU_{h,t}}{U_{h,t}}\right] - \frac{1}{2}\gamma_h\frac{1}{dt}E_t\left[\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right] + \frac{1}{dt}E_t\left[\frac{dU_{h,t}}{U_{h,t}}\frac{dM_{h,t}}{M_{h,t}}\right]\right).\end{aligned}$$

It follows from the above expression that the certainty equivalent operator, $\mu_{h,t}[\cdot]$, is given by

$$\begin{aligned}\frac{\mu_{h,t}[dU_{h,t}]}{dt} &= \frac{\mu_{h,t}[U_{h,t+dt}] - U_{h,t}}{dt} \\ &= U_{h,t}\left(\frac{1}{dt}E_t\left[\frac{dU_{h,t}}{U_{h,t}}\right] - \frac{1}{2}\gamma_h\frac{1}{dt}E_t\left[\left(\frac{dU_{h,t}}{U_{h,t}}\right)^2\right]\right).\end{aligned}$$

Therefore

$$\frac{\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[dU_{h,t}]}{dt} = \frac{\mu_{h,t}[dU_{h,t}]}{dt} + \frac{1}{dt}U_{h,t}E_t\left[\frac{dU_{h,t}}{U_{h,t}}\frac{dM_{h,t}}{M_{h,t}}\right],$$

i.e.

$$\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[dU_{h,t}] = \mu_{h,t}[dU_{h,t}] + U_{h,t}E_t\left[\frac{dU_{h,t}}{U_{h,t}}\frac{dM_{h,t}}{M_{h,t}}\right],$$

which implies

$$\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = \mu_{h,t}[U_{h,t+dt}] + U_{h,t}E_t\left[\frac{dU_{h,t}}{U_{h,t}}\frac{dM_{h,t}}{M_{h,t}}\right].$$

If the investment opportunity set is constant, then $U_{h,t}$ is a function solely of $W_{h,t}$ – there are no other state variables. Hence

$$E_t\left[\frac{dU_{h,t}}{U_{h,t}}\frac{dM_{h,t}}{M_{h,t}}\right] = \frac{W_{h,t}U_{W_{h,t}}}{U_{h,t}}\boldsymbol{\nu}^{\top}_{h,t}\boldsymbol{\omega}_{h,t}dt,$$

and so

$$\widehat{\mu}^{\boldsymbol{\nu}}_{h,t}[U_{h,t+dt}] = \mu_{h,t}[U_{h,t+dt}] + U_{h,t}\frac{W_{h,t}U_{W_{h,t}}}{U_{h,t}}\boldsymbol{\nu}^{\top}_{h,t}\boldsymbol{\omega}_{h,t}dt.$$

Equation (10) follows from the above expression and (8). □

Proof of Corollary 1

If the investment opportunity set is constant, then the personal divergence vector is obtained from the following optimization problem:

$$\inf_{\boldsymbol{\nu}_{h,t}} \left(\frac{W_{h,t} U_{W_{h,t}}}{U_{h,t}} \boldsymbol{\nu}_{h,t}^{\top} \boldsymbol{\omega}_{h,t} + \frac{1}{2\gamma_h} \sum_{n=1}^{N} \frac{\phi_{hn}}{1-\phi_{hn}} \frac{\nu_{hn,t}^2}{\sigma_n^2} \right).$$

The FOC for the above problem is

$$0 = \frac{W_{h,t} U_{W_{h,t}}}{U_{h,t}} \omega_{hn,t} + \frac{1}{\gamma_h} \frac{\phi_{hn}}{1-\phi_{hn}} \frac{\nu_{hn,t}}{\sigma_n^2}.$$

Equation (12) follows from the above equation. □

Proposition A1 and its proof

Proposition A1. *The utility function of a household with endogenous beliefs is given by the following Hamilton-Jacobi-Bellman equation:*

$$0 = \sup_{C_{h,t}} \left(\delta\, u_{\psi} \left(\frac{C_{ht}}{U_{ht}} \right) + \sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} \frac{1}{U_{h,t}} \mu_{h,t}^{\boldsymbol{\nu}} \left[\frac{dU_{h,t}}{dt} \right] \right), \quad \textit{where} \qquad \text{(A1)}$$

$$u_{\psi}(x) = \frac{x^{1-\frac{1}{\psi}} - 1}{1-\frac{1}{\psi}}, \psi > 0, \quad \textit{and}$$

$$\mu_{h,t}^{\boldsymbol{\nu}} \left[dU_{h,t} \right] = \mu_{h,t}^{\boldsymbol{\nu}} \left[U_{h,t+dt} - U_{h,t} \right] = \mu_{h,t}^{\boldsymbol{\nu}} \left[U_{h,t+dt} \right] - U_{h,t},$$

with $\mu_{h,t}^{\boldsymbol{\nu}} \left[U_{h,t+dt} \right]$ *given in* (10).

Proof: Writing out (14) explicitly gives

$$U_{h,t}^{1-\frac{1}{\psi_h}} = (1 - e^{-\delta_h dt}) C_{h,t}^{1-\frac{1}{\psi_h}} + e^{-\delta_h dt} \left(\mu_{h,t}^{\boldsymbol{\nu}} [U_{h,t+dt}] \right)^{1-\frac{1}{\psi_h}},$$

where for ease of notation sup and inf have been suppressed. Now

$$\begin{aligned}
\left(\mu_{h,t}^{\boldsymbol{\nu}} [U_{h,t+dt}] \right)^{1-\frac{1}{\psi_h}} &= \left(U_{h,t} + \mu_{h,t}^{\boldsymbol{\nu}} [dU_{h,t}] \right)^{1-\frac{1}{\psi_h}} \\
&= U_{h,t}^{1-\frac{1}{\psi_h}} \left(1 + \mu_{h,t}^{\boldsymbol{\nu}} \left[\frac{dU_{h,t}}{U_{h,t}} \right] \right)^{1-\frac{1}{\psi_h}} \\
&= U_{h,t}^{1-\frac{1}{\psi_h}} \left(1 + \left(1 - \frac{1}{\psi_h} \right) \mu_{h,t}^{\boldsymbol{\nu}} \left[\frac{dU_{h,t}}{U_{h,t}} \right] \right) + o(dt).
\end{aligned}$$

Hence

$$U_{h,t}^{1-\frac{1}{\psi_h}} = \delta C_{h,t}^{1-\frac{1}{\psi_h}} dt + U_{h,t}^{1-\frac{1}{\psi_h}} \left(1 + \left(1 - \frac{1}{\psi_h} \right) \mu_{h,t}^{\boldsymbol{\nu}} \left[\frac{dU_{h,t}}{U_{h,t}} \right] \right) - \delta_h U_{h,t}^{1-\frac{1}{\psi_h}} dt + o(dt),$$

from which we obtain (A1). □

From (A1), we see that the Hamilton-Jacobi-Bellman equation can be decomposed into a portfolio-optimization problem and an intertemporal consumption-choice problem. Given that the investment opportunity set that is constant over time, the maximized household utility is a constant multiple of the household's wealth. In this case, the Hamilton-Jacobi-Bellman equation can be decomposed into two parts, as shown in the proposition below.

Proposition A2 and its proof

Proposition A2. *The household's optimization problem consists of two parts, a single-period mean-variance optimization*

$$\sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} MV_h(\boldsymbol{\omega}_{h,t}, \boldsymbol{\nu}_{h,t}),$$

and an intertemporal consumption-choice problem

$$0 = \sup_{C_{h,t}} \left(\delta_h \, u_\psi \left(\frac{C_{h,t}}{U_{h,t}} \right) - \frac{C_{h,t}}{W_{h,t}} + \sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} MV_h(\boldsymbol{\omega}_{h,t}, \boldsymbol{\nu}_{h,t}) \right), \quad \textit{where} \tag{A2}$$

$$MV_h(\boldsymbol{\omega}_{h,t}, \boldsymbol{\nu}_{h,t}) = i + \big(\boldsymbol{\alpha} - i\mathbf{1}\big)^\top \boldsymbol{\omega}_{h,t} - \frac{1}{2}\gamma \boldsymbol{\omega}_{h,t}^\top V \boldsymbol{\omega}_{h,t} + \boldsymbol{\nu}_{h,t}^\top \boldsymbol{\omega}_{h,t} + \frac{1}{2\gamma_h} \boldsymbol{\nu}_{h,t}^\top \Sigma F_h^{-1} \Sigma \boldsymbol{\nu}_{h,t}, \tag{A3}$$

$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^\top$, *and* $\mathbf{1}$ *denotes the* $N \times 1$ *unit vector.*

Proof:

From (10) in Proposition 1, we have

$$\mu_{h,t}^{\boldsymbol{\nu}} \left[\frac{dU_{h,t}}{U_{h,t}} \right] = E_t \left[\frac{dU_{h,t}}{U_{h,t}} \right] - \frac{1}{2}\gamma_h E_t \left[\left(\frac{dU_{h,t}}{U_{h,t}} \right)^2 \right] + \left(\frac{W_{h,t} U_{W_{h,t}}}{U_{h,t}} \boldsymbol{\nu}_{h,t}^\top \boldsymbol{\omega}_{h,t} + L_{h,t} \right) dt$$

We use the Ansatz

$$U_{h,t} = u_h W_{h,t}, \tag{A4}$$

where u_h is a constant. Consequently,

$$\frac{dU_{h,t}}{U_{h,t}} = \frac{dW_{h,t}}{W_{h,t}},$$

and so.

$$\mu_{h,t}^{\boldsymbol{\nu}} \left[\frac{dU_{h,t}}{U_{h,t}} \right] = E_t \left[\frac{dW_{h,t}}{W_{h,t}} \right] - \frac{1}{2}\gamma_h E_t \left[\left(\frac{dW_{h,t}}{W_{h,t}} \right)^2 \right] + \left(\boldsymbol{\nu}_{h,t}^\top \boldsymbol{\omega}_{h,t} + L_{h,t} \right) dt.$$

Hence,

$$\frac{1}{dt}\mu^{\boldsymbol{\nu}}_{h,t}\left[\frac{dU_{h,t}}{U_{h,t}}\right] = -\frac{C_{h,t}}{W_{h,t}} + i + (\boldsymbol{\alpha} - i_t\mathbf{1})^\top \boldsymbol{\omega}_{h,t} - \frac{1}{2}\gamma_h \boldsymbol{\omega}_{h,t}^\top V \boldsymbol{\omega}_{h,t} + \boldsymbol{\nu}_{h,t}^\top \boldsymbol{\omega}_{h,t} + L_{h,t}$$

We now define

$$MV_h(\boldsymbol{\omega}_{h,t}, \boldsymbol{\nu}_{h,t}) = i + (\boldsymbol{\alpha} - i_t\mathbf{1})^\top \boldsymbol{\omega}_{h,t} - \frac{1}{2}\gamma_h \boldsymbol{\omega}_{h,t}^\top V \boldsymbol{\omega}_{h,t} + \boldsymbol{\nu}_{h,t}^\top \boldsymbol{\omega}_{h,t} + L_{h,t},$$

which can be rewritten as (A3). Therefore, provided the Ansatz in (A4) is true, our result follows from (A1) in Proposition A1.

We show the Ansatz in (A4) is true by explicitly solving for u_h. The first part of this proof consists of solving jointly for household beliefs and portfolios, which is done in the Proof of Proposition 2. The second part consists of solving for optimal consumption and substituting the optimal controls into (A2) and hence solving for $U_{h,t}$, which is also given in the proof of Proposition 2. □

Proof of Proposition 2

Substituting (A4) into (12) and exploiting the fact that Σ and F_h are diagonal matrices and hence commute with each other, we obtain

$$\boldsymbol{\nu}_{h,t} = -\gamma_h \Sigma^2 F_h^{-1}\, \boldsymbol{\omega}_{h,t}. \tag{A5}$$

Again, exploiting the commutativity of diagonal matrices, we can write the portfolio choice problem in (A3) as

$$\sup_{\boldsymbol{\omega}_{h,t}} (\boldsymbol{\alpha} - i\mathbf{1})^\top \boldsymbol{\omega}_{h,t} - \frac{1}{2}\gamma_h \boldsymbol{\omega}_{h,t}^\top \left(V + \Sigma^2 F_h^{-1}\right) \boldsymbol{\omega}_{h,t}.$$

The above linear-quadratic problem has a unique interior solution given by (16). Substituting (16) into (A5) and simplifying gives (15). We can also rewrite the expression for $\boldsymbol{\omega}_h$ in (16) in terms of the personal divergence measure:

$$\boldsymbol{\omega}_h = \frac{1}{\gamma_h} V^{-1}(\boldsymbol{\alpha} + \boldsymbol{\nu}_h - i\mathbf{1}),$$

where

$$\boldsymbol{\nu}_h = -(I + V F_h \Sigma^{-2})(\boldsymbol{\alpha} - i\mathbf{1}).$$

Substituting the optimal controls (15) and (16) into (A3), and simplifying gives

$$\sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} MV_h(\boldsymbol{\omega}_{h,t}, \boldsymbol{\nu}_{h,t}) = i + \frac{1}{2\gamma_h}(\boldsymbol{\alpha} - i\mathbf{1})^\top (V + \Sigma^2 F_h^{-1})^{-1} (\boldsymbol{\alpha} - i\mathbf{1}).$$

From the Hamilton-Jacobi-Bellman equation in (A1) (and of course also the specialized version in (A2)), the first-order condition with respect to consumption is

$$\delta_h \left(\frac{C_{h,t}}{W_{h,t}}\right)^{-\frac{1}{\psi_h}} = \frac{U_{h,t}}{W_{h,t}}. \tag{A6}$$

From the above expression (which holds also for a stochastic investment opportunity set), we can see that $\frac{U_{h,t}}{W_{h,t}}$ is a constant if and only if $\frac{C_{h,t}}{W_{h,t}}$ is a constant. Substituting the above first-order condition into (A2) to eliminate $U_{h,t}$ and then solving for $\frac{C_{h,t}}{W_{h,t}}$ gives

$$c_h = \frac{C_{h,t}}{W_{h,t}} = \delta_h \psi_h + (1-\psi_h) \sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} MV_h(\boldsymbol{\omega}_{h,t}, \boldsymbol{\nu}_{h,t}).$$

We can thus see that $\frac{C_{h,t}}{W_{h,t}}$ is indeed a constant. Also, from (A6) we can see that

$$u_h = \left[\frac{\delta_h^{\psi_h}}{c_h}\right]^{\frac{1}{\psi_h - 1}}.$$

Therefore, we obtain the following result for a household's maximised utility

$$u_h = \frac{U_{h,t}}{W_{h,t}} = \left[\frac{\delta_h^{\psi_h}}{\delta_h \psi_h + (1-\psi_h) \sup_{\boldsymbol{\omega}_{h,t}} \inf_{\boldsymbol{\nu}_{h,t}} MV_h(\boldsymbol{\omega}_{h,t}, \boldsymbol{\nu}_{h,t})}\right]^{\frac{1}{\psi_h - 1}}.$$

□

B Quantfiying the Impact of Geography Beliefs

We want to show how geography impacts the deviation of beliefs relative to rational expectations. The reference probability measure $\mathbb{P}$ is the belief under rational expectations. The deviation of the non-rational expectations belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ from $\mathbb{P}$ is given by the relative entropy per unit time of $\mathbb{P}$ with respect $\mathbb{Q}^{\boldsymbol{\nu}_h}$, i.e. (6).

We do not directly observe the vector $\boldsymbol{\nu}_h$, which we need in order to compute (6). However, we can indirectly obtain $\boldsymbol{\nu}_h$ via knowledge of portfolios, as shown in (12) (and more simply in (A5)). We now eliminate $\boldsymbol{\nu}_h$ in (6) by substituting (A5) into (6), thereby obtaining

$$D^{KL}[\mathbb{P}|\mathbb{Q}^{\boldsymbol{\nu}_h}] = \gamma_h^2 \boldsymbol{\omega}_{h,t}^\top \Sigma F_h^{-1} (\Omega^{-1})^\top F_h^{-1} \Sigma \boldsymbol{\omega}_{h,t}. \tag{B1}$$

The above expression allows us to measure the deviation of the non-rational expectations belief, $\mathbb{Q}^{\boldsymbol{\nu}_h}$ from the rational expectations belief $\mathbb{P}$ using portfolio data.

To understand (B1), we evaluate it for the special case when the return correlations are zero, i.e. $\Omega = I$. Therefore, (B1) reduces to

$$\begin{aligned} D^{KL}[\mathbb{P}|\mathbb{Q}^{\boldsymbol{\nu}_h}] &= \gamma_h^2 \boldsymbol{\omega}_{h,t}^\top \Sigma F_h^{-2} \Sigma \boldsymbol{\omega}_{h,t} \\ &= \sum_{n=1}^{N} \omega_{hn,t}^2 \sigma_n^2 \left(\frac{1}{\phi_{hn}} - 1 \right)^2 . \end{aligned} \tag{B2}$$

With portfolio weights given by

$$\omega_{hn} = \frac{1}{\gamma_h} \frac{\alpha_n - i}{\sigma_n^2} \phi_{hn},$$

the deviation of the non-rational expectations belief, $\mathbb{Q}^{\boldsymbol{\nu}_h}$ from the rational expectations belief $\mathbb{P}$ becomes

$$D^{KL}[\mathbb{P}|\mathbb{Q}^{\boldsymbol{\nu}_h}] = \sum_{n=1}^{N} \left(\frac{\alpha_n - i}{\sigma_n} \right)^2 (1 - \phi_{hn})^2 , \tag{B3}$$

which is decreasing in each individual proximity measure, ϕ_{hn}.

C Welfare Analysis

We measure household-level welfare losses by taking the optimal controls for a household which has biased beliefs and substituting them into the utility of a household whose beliefs are given by the reference probability measure $\mathbb{P}$.

For an exogenously specified consumption and portfolio rule, (A2) reduces to

$$0 = \delta_h \, u_{\psi_h} \left(\frac{C_{h,t}}{U_{h,t}} \right) - \frac{C_{h,t}}{W_{h,t}} + U_h^{MV}(\boldsymbol{\omega}_{h,t}), \tag{C1}$$

where $U_h^{MV}(\boldsymbol{\omega}_{h,t})$ is the utility of a mean-variance investor whose beliefs are given by the reference probability measure $\mathbb{P}$, i.e.

$$U_h^{MV}(\boldsymbol{\omega}_{h,t}) = i + \left(\boldsymbol{\alpha} - i\mathbf{1} \right)^\top \boldsymbol{\omega}_{h,t} - \frac{1}{2} \gamma \boldsymbol{\omega}_{h,t}^\top V \boldsymbol{\omega}_{h,t}.$$

By making $U_{h,t}$ the subject of (C1), we see that for a for a household whose beliefs are given by the reference probability measure $\mathbb{P}$, utility per unit wealth, v_h is given in terms of consumption-portfolio choices by

$$v_h(c_h, \boldsymbol{\omega}_h) = \left[\frac{\psi_h \delta_h}{\psi_h \delta_h + (1 - \psi_h)(U_h^{MV}(\boldsymbol{\omega}_h) - c_h)} \right]^{\frac{1}{1 - 1/\psi_h}} c_h,$$

where $c_h = C_h / W_h$ is the consumption-wealth ratio.

C.1 Parameter values

- We already have an estimate for κ, so we can calculate the matrix F_h for each household.
- In addition, we estimate from the data individual stock-return volatilities, cross-correlations and expected excess returns.
- Finally, we need to assume values for the preference parameters: δ_h, ψ_h and γ_h, which we can assume are equal across households (and hence, across post codes), which will allow us to focus on differences arising purely from location. Given that there is some debate about the appropriate values for these parameters, we can consider three values for each of these parameters: a low value, a medium value, and a high value.
  - $\delta_h = \{0.02, 0.03, 0.04\}$.
  - $\psi_h = \{0.70, 0.80, 0.90\}$.
  - $\gamma_h = \{2.00, 4.00, 6.00\}$.

C.2 Measuring Beliefs & Welfare

1. Using the variance-covariance matrix for returns, we compute, at the post code level, each household's portfolio volatility $\sqrt{\boldsymbol{\omega}_{h,t}^{\top} V \boldsymbol{\omega}_{h,t}}$. The advantage of this measure is that it does not depend on expected excess returns, which are notorious for being difficult to measure, and it is independent of preference parameters.

2. Using, at the post code level, each household's portfolio vector, we compute the deviation of the implied non-rational expectations belief $\mathbb{Q}^{\boldsymbol{\nu}_h}$ from the rational expectations belief $\mathbb{P}$, via (B2). Note that this quantity is also a measure of information loss per unit time, and is decreasing in each of the proximity parameters ϕ_{hn}, as we can see from (B3).

   In order to make the measure (B2) independent of the value chosen for the preference parameter γ_h (which we assume is equal across households), we could report a scaled version of $D^{KL}[\mathbb{P}|\mathbb{Q}^{\boldsymbol{\nu}_h}]$; for instance, $\frac{D^{KL}[\mathbb{P}|\mathbb{Q}^{\boldsymbol{\nu}_h}]}{\sum_{k=1}^{H} D^{KL}[\mathbb{P}|\mathbb{Q}^{\boldsymbol{\nu}_k}]}$. We could then produce a heatmap of Finland reporting the relative information loss for each post code, as a way of measuring the effect of geography on the deviation from the rational expectations belief, using a darker shade of red for larger deviations.

3. Using, at the post code level, each household's portfolio vector, we compute its network-weighted information losses per unit time,

$$\hat{L}_{h,t} = \gamma_h^2 \boldsymbol{\omega}_{h,t}^\top \Sigma^2 F_h^{-1} \boldsymbol{\omega}_{h,t}.$$

Observe that the above expression can be rewritten as

$$\hat{L}_{h,t} = \gamma_h^2 \sum_{n=1}^{N} \omega_{hn,t}^2 \sigma_n^2 \left(\frac{1}{\phi_{hn}} - 1 \right). \tag{C2}$$

Since ϕ_{hn} is a measure of proximity, we see that network-weighted information losses shall be larger when proximity is smaller.

For an example, consider also the case where $N = 2$ and $\phi_{h1} = 1$, $\phi_{h2} = \epsilon$ (where ϵ is small and positive), and $\sigma_1 = \sigma_2 = \sigma$. The first asset is the 'close' one and the second one is the 'far' asset. In this case,

$$\hat{L}_{h,t} = \gamma_h^2 \sigma^2 \omega_{h2,t}^2 \left(\frac{1}{\epsilon} - 1 \right).$$

We can see that the household incurs greater network-weighted information losses when she invests more of her wealth in the far asset. She will trade off this cost against any benefit from investing in the far asset. With $\phi_{h1} = \phi_{h2} = \phi$, we would obtain

$$\hat{L}_{h,t} = \gamma_h^2 \sigma^2 \left(\frac{1}{\phi} - 1 \right) (\omega_{h1,t}^2 + \omega_{h2,t}^2).$$

This case is the standard case of global, i.e. non-asset-specific, ambiguity aversion.

Now we substitute the portfolio policy $\omega_{hn} = \frac{1}{\gamma_h} \frac{\alpha_n - i}{\sigma_n^2} \phi_{hn}$ into (C2) to obtain

$$\hat{L}_{h,t} = \sum_{n=1}^{N} \left(\frac{\alpha_n - i}{\sigma_n} \right)^2 \phi_{hn} \left(1 - \phi_{hn}\right).$$

In contrast with (B3), the above expression is not monotonic in ϕ_{hn}. This is a consequence of the network weighting of the information losses.

4. We can compute a household's mean-variance utility,

$$U_{h,t}^{MV} = i_t + (\boldsymbol{\alpha} - i\mathbf{1})^\top \boldsymbol{\omega}_{h,t} - \frac{1}{2} \gamma_h \boldsymbol{\omega}_{h,t}^\top V \boldsymbol{\omega}_{h,t}, \tag{C3}$$

which will depend on the risk-aversion parameter, γ_h. So, for each post code, we would report the mean-variance utility for three levels of γ_h. We could then produce a heat map using the middle value of the risk aversion parameter.

For the special case where return correlations are zero and return volatilities and expected returns are equal across assets, we know $V = \sigma^2 I$ and $\omega_{hn} = \frac{1}{\gamma_h} \frac{\alpha - i}{\sigma^2} \phi_{hn}$. In this case, (C3) reduces to

$$U_{h,t}^{MV} = i_t + \frac{1}{2\gamma_h} \left(\frac{\alpha - i}{\sigma} \right)^2 \sum_{n=1}^{N} \phi_{hn},$$

which increases with respect to each proximity measure ϕ_{hn}.

5. We could also compute each household's consumption-wealth policy,

$$c_h = \frac{C_{h,t}}{W_{h,t}} = \psi_h \delta_h + (1 - \psi_h) \left(i_t + \frac{1}{2} \gamma_h \boldsymbol{\omega}_{h,t}^{\top} V \boldsymbol{\omega}_{h,t} + \frac{1}{2\gamma_h} \hat{L}_{h,t} \right). \tag{C4}$$

This depends on all three preference parameters, so we would have nine values for each household, corresponding to the three values for the three preference parameters.

For the special case where return correlations are zero and return volatilities and expected returns are equal across assets, we know $V = \sigma^2 I$ and $\omega_{hn} = \frac{1}{\gamma_h} \frac{\alpha - i}{\sigma^2} \phi_{hn}$. In this case (C4) reduces to

$$\begin{aligned} c_h &= \frac{C_{h,t}}{W_{h,t}} \\ &= \psi_h \delta_h + (1 - \psi_h) \left(i_t + \frac{1}{2\gamma_h} \left(\frac{\alpha - i}{\sigma} \right)^2 \sum_{n=1}^{N} \phi_{hn} + \frac{1}{2\gamma_h} \sum_{n=1}^{N} \left(\frac{\alpha - i}{\sigma} \right)^2 \phi_{hn} \left(1 - \phi_{hn}\right) \right) \\ &= \psi_h \delta_h + (1 - \psi_h) \left(i_t + \frac{1}{2\gamma_h} \left(\frac{\alpha - i}{\sigma} \right)^2 \left(2 \sum_{n=1}^{N} \phi_{hn} - \sum_{n=1}^{N} \phi_{hn}^2 \right) \right) \\ &= \psi_h \delta_h + (1 - \psi_h) \left(i_t + \frac{1}{2\gamma_h} \left(\frac{\alpha - i}{\sigma} \right)^2 \sum_{n=1}^{N} \left[1 - (1 - \phi_{hn})^2 \right] \right). \end{aligned}$$

The above expression is useful, because it tells us how biases distort the consumption-savings decision. For $\psi_h < 1$ ($\psi_h > 1$), decreasing a particular ϕ_{hn} leads to decreased (increased) consumption from wealth.

6. Finally, we can compute each household's lifetime welfare per unit wealth.

$$\frac{U_{h,t}}{W_{h,t}} = \left[\frac{\delta_h \psi_h}{\psi_h \delta_h + (1 - \psi_h)(U_{h,t}^{MV} - c_{h,t})} \right]^{\frac{1}{1 - 1/\psi_h}} c_h,$$

which also depends on all three preference parameters, so we would have nine values for each household. We could then produce a heat map using the middle values of the three preference parameters.

To understand the above expression, observe that increasing current consumption is clearly beneficial but comes at the expense of reduced savings. The benefit is represented by the last factor i.e., the c_h that appears outside the expression in square brackets. The negative impact on savings is captured by the term $U^{MV}_{h,t} - c_{h,t}$ in the denominator of the first component. The term $U^{MV}_{h,t} - c_{h,t}$ is the risk-adjusted expected return on a household's wealth, net of consumption. To obtain its impact on lifetime welfare, this expected return needs to be capitalized, as shown above. The capitalized value depends on the intertemporal aspects of the household's preferences, that is, the rate of time preference, δ_h, and the elasticity of intertemporal substitution, ψ_h.[14]

[14] In our partial equilibrium setting, U_h/W_h is non-monotonic in ϕ_{hn}. If households had identical preferences, imposing market clearing would give an endogenous risk-free rate and the U_h/W_h would simplify to give an expression which is monotonic in ϕ_{hn}.

D Additional Details for the Empirical Results

In this appendix, we provide additional information about the empirical tests and results.

Bootstrapping Method

Consider a finite multi-set of vectors, $\mathcal{V} = \{\mathbf{v}_m\}_{m\in\mathcal{M}}$, $|\mathcal{M}| \in \mathbb{N}$, where $\mathbf{v}_m \in \{0,1\}^N$, $N \in \mathbb{N}$. In our test, m will be associated with a household and $(\mathbf{v}_m)_n$ will represent whether it holds stock n (in which case $(\mathbf{v}_m)_n = 1$) or not (in which case $(\mathbf{v}_m)_n = 0$). The object $\mathcal{V}$ is a multi-set, since there may be multiple households with the same stocks in their portfolios. The nonzero elements in $\mathbf{v} \in \mathcal{V}$ are denoted the vector's *constituents*, and $c_{\mathbf{v}} = |\{\ell : (\mathbf{v})_\ell = 1\}|$ is defined as the number of such constituents.

For two elements, $\mathbf{v}^a \in \mathcal{V}$, $\mathbf{v}^b \in \mathcal{V}$, define the vector of overlaps, $\mathcal{O}(\mathbf{v}^a, \mathbf{v}^b) \in \{0,1\}^N$, such that $\mathcal{O}(\mathbf{v}^a, \mathbf{v}^b)_n = (\mathbf{v}^a)_n \times (\mathbf{v}^b)_n$, and the overlap $o(\mathbf{v}^a, \mathbf{v}^b) = \sum_n \mathcal{O}(\mathbf{v}^a, \mathbf{v}^b)_n \in \{0, 1, \ldots, N\}$.

Consider a multi-set $\mathcal{A}$ of (at least two) elements in $\mathcal{V}$, ordered from 1 to $r = |\mathcal{A}|$, $\mathbf{a}_1, \ldots, \mathbf{a}_r$. This set will correspond to the set of household portfolios in a postal code area in our test. The overlap in $\mathcal{A}$ is

$$o(\mathcal{A}) = \frac{1}{2} \sum_{\substack{k,\ell=1,\ldots,r \\ k\neq\ell}} o(\mathbf{a}_k, \mathbf{a}_\ell). \tag{D1}$$

Define $\mathcal{Z}^k = \{\mathbf{v} \in \mathcal{V} : c_{\mathbf{v}} = k\}$, the set of all elements in $\mathcal{V}$, with k constituents, and $z^k = |\mathcal{Z}^k|$, the number of such elements. Also define $K = \max\{k : z^k > 0\}$. For simplicity, we assume that $z^k > 0$ for all $k < K$. Our analysis also holds if this is assumption is not satisfied, but with the extra notational burden of excluding such k's for which $z^k = 0$ in some of the sums below, which is why we make this assumption.

Consider a random data generating process to generate a vector $\tilde{\mathbf{a}}$ of k constituents from the elements in $\mathcal{Z}^k$, where each element in $\mathcal{Z}^k$ is chosen with the same probability. Thus, if the elements in $\mathcal{Z}^k$ are ordered as $\mathbf{v}_1^k, \ldots, \mathbf{v}_{z_k}^k$, then

$$\mathbb{P}(\tilde{\mathbf{a}} = \mathbf{v}_\ell^k) = \frac{1}{z_k}, \qquad \ell = 1, \ldots, k.$$

Now, consider two randomly chosen such elements, $\tilde{\mathbf{a}} \in \mathcal{Z}^k$ and $\tilde{\mathbf{b}} \in \mathcal{Z}^\ell$, $1 \leq k, \ell \leq K$. The expected overlap of these two elements, given the above data generating process, is then

$$o_{k,\ell}^{\text{Exp}} = E\left[o(\tilde{\mathbf{a}}, \tilde{\mathbf{b}})\right] = \frac{1}{z^k z^\ell} \sum_{s=1}^{z_k} \sum_{t=1}^{z_\ell} o(\mathbf{v}_s^k, \mathbf{v}_t^\ell).$$

Moreover, given the set $\mathcal{A}$ with $r \geq 2$ elements, and the ordered elements, $\mathbf{a}_1, \ldots, \mathbf{a}_r$, the expected overlap, when $\mathcal{A}$ is chosen from the above data generating process, is

$$o^{\text{Exp}}(\mathcal{A}) = \frac{1}{2} \sum_{\substack{k,\ell=1,\ldots,r \\ k \neq \ell}} o^{\text{Exp}}_{c_{\mathbf{a}_k}, c_{\mathbf{a}_\ell}}. \tag{D2}$$

By comparing the actual overlap given in equation (D1) with the expected overlap given in equation (D2), we can thus draw inferences about whether the elements in $\mathcal{A}$ were randomly selected according to the above data generating process, or whether they have too much (or too little overlap) compared with what would be expected.

Note that it is straightforward to restrict the above test to portfolios up to a specific size, S, by replacing the set $\mathcal{A}$ with $\mathcal{A}^S = \{\mathbf{v} \in \mathcal{A} : c_{\mathbf{v}} \leq S\}$. This allows us to draw additional inferences about whether potential nonrandomness stems from elements with a low number of constituents or not. Of course, $\mathcal{A}^N = \mathcal{A}$.

For the portfolio data, we use the set of portfolios of households in different postal code areas, $\mathcal{A}_p = \{\mathbf{v}_h : h \in \mathcal{H}_p\}$, $p = 1, \ldots, P$. Our null hypothesis is that portfolios of households within a postal code are randomly selected from the total set of portfolios, $\mathcal{V} = \cup_p \mathcal{A}_p$. We compare the actual and expected total overlap across postal codes areas, controlling for portfolio size, i.e.,

$$N^c_{TOT} = \sum_p o(\mathcal{A}^S_p) \qquad \text{with} \qquad E\left[\tilde{N}^c_{TOT}\right] = \sum_p o^{\text{Exp}}(\mathcal{A}^S_p),$$

where we vary the portfolio size, S, between 1 and N.

The advantage of this method, which is admittedly more complex than the simple test in the main paper, is that instead of assuming that each household picks each stock with probability $\frac{1}{N}$, we use the empirical distribution function of portfolios to test whether the actual portfolios observed within postal codes are consistent with a "random allocation" of these portfolios. This approach is therefore robust to some stocks being much more widely held than others. Note that our total test contains $P = 2,923$ postal codes, with an average of 139 households each. The overrepresentation of overlap of 43.6%-139% we find is therefore extremely statistically significant.

Additional Tables

The stocks in our sample are shown in Table D1. The estimate sensitivity coefficient when excluding the Helsinki area is shown in Table D2. The estimated sensitivity coefficient when including only distant observations are shown in Table D3.

Table D1: Stocks in sample

This table lists the 125 stocks issued by companies that are headquartered in Finland and are included in our sample. For companies with A- and B-shares, both shares are included in our sample.

1	Bank of Aland Plc A	2	Pohjola Group Plc
3	Norvestia Plc	4	Kesko Corporation B
5	Stockmann Plc A	6	Stockmann Plc B
7	Tieto Corporation	8	Amer Sports Corporation
9	Fiskars Corporation	10	Fiskars Corporation K
11	Huhtamki Oyj	12	Instrumentarium
13	Kone Corporation B	14	Metsa Board Oyj A
15	Metsa Board Oyj B	16	Nokia Corporation
17	Tamro Oyj	18	Tamfelt Corp.
19	Tamfelt Corporation Ord. shares	20	Bank of Aland Plc B
21	Uponor Oyj	22	Outokumpu Oyj
23	Citycon Oyj	24	Polar Real Estate Corp.
25	Raisio Plc Vaihto-osake	26	Birka Line Abp B
27	Pohjola Bank A	28	Finnair Oyj
29	Sampo Plc A	30	Stromsdal Corporation
31	Apetit Plc	32	Rautaruukki Corporation
33	Finnlines Plc	34	Silja Oyj Abp
35	Wartsila Corporation A	36	Wartsila Corporation
37	Tiimari Plc	38	Kemira Oyj
39	Ponsse Oyj	40	Viking Line Abp
41	Nokian Tyres Plc	42	Biohit Oyj B
43	Konecranes Plc	44	Stora Enso Oyj A
45	Stora Enso Oyj R	46	UPM-Kymmene Corporation
47	HKScan Oyj A	48	PKC Group Oyj
49	Incap Corporation	50	Atria Plc A
51	Payry PLC	52	Sponda Plc
53	Technopolis Plc	54	Valoe Oyj
55	Alma Media Corporation 1	56	Alma Media Corporation 2
57	Ramirent Plc	58	Fortum Corporation
59	Bittium Corporation	60	Yomi Plc
61	Rapala VMC Corporation	62	Sonera Oyj
63	Eimo Oyj	64	Innofactor Plc
65	Marimekko Corporation	66	SanomaWSOY Corporation A
67	Sanoma Corporation	68	Teleste Corporation
69	Oral Hammaslakarit Plc	70	Perlos Corporation
71	Metso Corporation	72	Talentum Oyj
73	Kesko Corporation A	74	Aldata Solution Oyj
75	Digia Plc	76	Solteq Oyj
77	Ixonos Plc	78	Aspo Plc
79	Aspocomp Group Plc	80	Dovre Group Plc
81	Trainers House Plc	82	Comptel Corporation
83	SSH Communications Security Oyj	84	Basware Corporation
85	Wulff Group Plc	86	Saunalahti Group Oyj
87	Etteplan Oyj	88	QPR Software Plc
89	eQ Oyj	90	Tekla Corporation
91	Sievi Capital plc	92	Sentera Plc
93	Okmetic Oyj	94	CapMan Plc B
95	Vacon Plc	96	eQ Oyj
97	Evox Rifa Group Plc	98	Componenta Corporation
99	Glaston Corporation	100	Tecnotree Corporation
101	Lassila & Tikanoja Plc	102	Suominen Oyj
103	Revenio Group Corporation	104	Biotie Therapies Corp.
105	Ilkka-Yhtyma Oyj 2	106	Neo Industrial Oyj
107	Orion Corporation A	108	Orion Corporation B
109	Raisio Plc K	110	Saga Furs Oyj C
111	YIT Corporation	112	Stonesoft Corporation
113	F-Secure Corporation	114	Chips Corporation B
115	Efore Plc	116	Hackman Oyj Abp
117	Honkarakenne Oyj B	118	Lemminkainen Corporation
119	Evia Oyj	120	Martela Oyj A
121	Olvi Plc A	122	Cramo Oyj
123	Tulikivi Oyj A	124	Elecster Oyj A
125	Vaisala Corporation A		

Table D2: Estimate of sensitivity coefficient excluding Helsinki area

This table estimates the sensitivity coefficient, κ, using only firms and households outside of the Helsinki area (with 5-digit postal codes). Panel A includes observations with $\omega_{pn} = 0$, replacing with $\omega_{pn} = 1$ (one Finnish Mark, corresponding to approximately USD 0.17). Panel B excludes observations with $\omega_{pn} = 0$. Univariate in column 2, including risk aversion in column 3, including risk aversion and stock distributions in column 4, panel regression with postal code and stock fixed effects and robust standard errors double-clustered at the postal code and firm level in column 5. Statistical significance levels: * = 0.01, ** = 0.001, *** = 0.0001.

(1)	(2)	(3)	(4)	(5)
Panel A				
Sensitivity coefficient, κ	4.228***	3.868***	3.771***	3.771***
Standard error	0.069	0.071	0.073	0.630
log risk aversion, g				
-average		-3.806	-3.771	
-max		-1.097	-1.073	
-min		-12.393	-12.355	
log distribution, s				
-average			0	
-max			5.418	
-min			-1.797	
R^2	0.051	0.308	0.487	0.487
Adj. R^2	0.051	0.283	0.469	0.469
$N = 69,368$				
Panel B				
Sensitivity coefficient, κ	1.824***	1.892***	2.754***	2.754***
Standard error	0.074	0.082	0.071	0.325
log risk aversion, g				
-average		-8.072	-8.084	
-max		-2.681	-2.282	
-min		-12.392	-12.757	
log distribution, s				
-average			0	
-max			1.787	
-min			-3.197	
R^2	0.028	0.249	0.501	0.501
Adj. R^2	0.028	0.159	0.440	0.440
$N = 20,787$				

Table D3: Estimate of sensitivity coefficient including only distant observations

This table estimates the sensitivity coefficient, κ, including only postal code/firm observation at distances larger than d_0, and excluding $\omega_{pn} = 0$ observations. Panel A sets d_0 to 0.01, corresponding to a minimum distance between postal codes of account holder and firm headquarter of about 8 miles, whereas in Panel B $d_0 = 0.03$, corresponding to a minimal distance of about 24 miles. Univariate in column 2, including risk aversion in column 3, including risk aversion and stock distributions in column 4, panel regression with postal code and stock fixed effects and robust standard errors double-clustered at the postal code and firm level in column 5. Statistical significance levels: * = 0.01, ** = 0.001, *** = 0.0001.

(1)	(2)	(3)	(4)	(5)
Panel A				
Sensitivity coefficient, κ	1.772***	1.844***	2.397***	2.397***
Standard error	0.036	0.063	0.048	0.392
log risk aversion, g				
-average		-8.480	-7.899	
-max		-4.543	-2.282	
-min		-12.415	-12.609	
log distribution, s				
-average			0	
-max			4.142	
-min			-3.952	
R^2	0.018	0.186	0.614	0.614
Adj R^2	0.018	0.168	0.605	0.605
$N = 129,802$				
Panel B				
Sensitivity coefficient, κ	0.931***	0.991***	1.662***	1.662***
Standard error	0.039	0.070	0.053	0.326
log risk aversion, g				
-average		-8.187	-7.680	
-max		-4.000	-2.561	
-min		-11.798	-12.117	
log distribution, s				
-average			0	
-max			4.067	
-min			-3.977	
R^2	0.005	0.160	0.599	0.599
Adj. R^2	0.005	0.139	0.589	0.589
$N = 119,874$				

References

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2012, The network origins of aggregate fluctuations, *Econometrica* 80, 1977–2016.

Acemoglu, Daron, and Asu Ozdaglar, 2009, Networks, course 14.15J: Massachusetts Institute of Technology: MIT OpenCourseWare, https://ocw.mit.edu.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2015, Systemic risk and stability in financial networks, *American Economic Review* 105, 564–608.

Ahern, Kenneth R, 2013, Network centrality and the cross section of stock returns, SSRN Working Paper.

Ahern, Kenneth R, and Jarrad Harford, 2014, The importance of industry links in merger waves, *The Journal of Finance* 69, 527–576.

Allen, Franklin, and Ana Babus, 2009, Networks in finance, in Paul R. Kleindorder, Yoram Wind, and Robert E. Gunther, (eds.) *The Network Challenge: Strategy, Profit, and Risk in an Interlinked World* (Pearson).

Bailey, Michael, Rachel Cao, Theresa Kuchler, and Johannes Stroebel, 2018, The economic effects of social networks: Evidence from the housing market, forthcoming in *Journal of Political Economy.*

Baltzer, Markus, Oscar Stolper, and Andreas Walter, 2013, Is local bias a cross-border phenomenon? Evidence from individual investors' international asset allocation, *Journal of Banking & Finance* 37, 2823–2835.

Baltzer, Markus, Oscar Stolper, and Andreas Walter, 2015, Home-field advantage or a matter of ambiguity aversion? Local bias among german individual investors, *European Journal of Finance* 21, 734–754.

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho, 2019, Survey data and subjective beliefs in business cycle models, University of Minnesota Working Paper.

Bodnaruk, Andriy, 2009, Proximity always matters: Local bias when the set of local companies changes, *Review of Finance* 13, 629–656.

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio, 2011, Uncertainty averse preferences, *Journal of Economic Theory* 146, 1275–1330.

Chen, Zengjing, and Larry G. Epstein, 2002, Ambiguity, risk, and asset returns in continuous time, *Econometrica* 70, 1403–1443.

Chiang, Yao-Min, David Hirshleifer, Yiming Qian, and Ann E. Sherman, 2011, Do investors learn from experience? Evidence from frequent IPO investors, *The Review of Financial Studies* 24, 1560–1589.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick, 2009, Reinforcement learning and savings behavior, *The Journal of Finance* 64, 2515–2534.

Cohen, Lauren, 2009, Loyalty-based portfolio choice, *Review of Financial Studies* 22, 1213–1245.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *The Journal of Finance* 63, 1977–2011.

Coval, J., and T. Moskovitz, 1999a, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109, 811–841.

Coval, J., and T. Moskovitz, 1999b, Home bias at home: Local equity preference in domestic portfolios, *The Journal of Finance* 54, 2045–2073.

Cox, John C., Jr. Ingersoll, Jonathan E., and Stephen A. Ross, 1985, An intertemporal general equilibrium model of asset prices, *Econometrica* 53, 363–84.

Das, Sreyoshi, Camelia M. Kuhnen, and Stefan Nagel, 2017, Socioeconomic status and macroeconomic expectations, Working Paper, National Bureau of Economic Research.

Døskeland, Trond M, and Hans K Hvide, 2011, Do individual investors have asymmetric information based on work experience?, *The Journal of Finance* 66, 1011–1041.

Dow, James, and Sergio Ribeiro da Costa Werlang, 1992, Uncertainty aversion, risk aversion, and the optimal choice of portfolio, *Econometrica* 60, 197–204.

Duffie, J. Darrell, and Larry G. Epstein, 1992, Stochastic Differential Utility, *Econometrica* 60, 353–394.

Easley, David, and Jon Kleinberg, 2010, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World* (Cambridge University Press).

Epstein, Larry G., and Tan Wang, 1994, Intertemporal asset pricing under Knightian uncertainty, *Econometrica* 62, 283–322.

Epstein, Larry G., and Stanley Zin, 1989, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Farboodi, Maryam, 2017, Intermediation and voluntary exposure to counterparty risk, SSRN Working Paper.

Gabaix, Xavier, 2011, The granular origins of aggregate fluctuations, *Econometrica* 79, 733–772.

Gagliardini, Patrick, Paolo Porchia, and Fabio Trojani, 2008, Ambiguity aversion and the term structure of interest rates, *The review of financial studies* 22, 4157–4188.

Gilboa, Itzhak, and David Schmeidler, 1989, Maxmin expected utility theory with non-unique prior, *Journal of Mathematical Economics* 18, 141–153.

Greenwood, Robin, and Stefan Nagel, 2009, Inexperienced investors and bubbles, *Journal of Financial Economics* 93, 239–258.

Grinblatt, Mark, and Matti Keloharju, 2000, The investment behavior and performance of various investor types: A study of Finland's unique data set, *Journal of Financial Economics* 55, 43–57.

Grinblatt, Mark, and Matti Keloharju, 2001, How distance, language, and culture influence stockholdings and trades, *The Journal of Finance* 56, 1053–1073.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2006, Does culture affect economic outcomes?, *Journal of Economic Perspectives* 20, 23–48.

Hansen, Lars Peter, and Thomas J. Sargent, 2007, *Robustness* (Princeton University Press, Princeton, NJ).

Herskovic, Bernard, and Joao Ramos, 2017, Acquiring information through peers, UCLA Working Paper.

Huberman, Gur, 2001, Familiarity breeds investment, *The Review of Financial Studies* 14, 659–680.

Jackson, Matthew O., 2010, *Social and Economic Networks* (Princeton University Press).

Jackson, Matthew O, 2014, Networks in the understanding of economic behaviors, *Journal of Economic Perspectives* 28, 3–22.

Kaustia, Markku, and Samuli Knüpfer, 2008, Do investors overweight personal experience? Evidence from IPO subscriptions, *The Journal of Finance* 63, 2679–2702.

Kaustia, Markku, and Samuli Knüpfer, 2012, Peer performance and stock market entry, *Journal of Financial Economics* 104, 321–338.

Knüpfer, Samuli, Elias Rantapuska, and Matti Sarvimäki, 2017, Formative experiences and portfolio choice: Evidence from the Finnish Great Depression, *The Journal of Finance* 72, 133–166.

Kuchler, Theresa, and Basit Zafar, 2018, Personal experiences and expectations about aggregate outcomes, Working Paper, New York University, Stern School of Business.

Landier, Augustin, Yueran Ma, and David Thesmar, 2017, New experimental evidence on expectations formation, Working Paper, MIT.

Laudenbach, Christine, Ulrike Malmendier, and Alexandra Niessen-Ruenzi, 2018, The long-lasting effects of propaganda on financial risk-taking, Working Paper.

Lindblom, Ted, Taylan Mavruk, and Stefan Sjögren, 2018, East or west, home is best: The birthplace bias of individual investors, *Journal of Banking & Finance* 92, 323–339.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, 2006, Ambiguity aversion, malevolent nature, and the variational representation of preferences, *Econometrica* 74, 1447–1498.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression babies: Do macroeconomic experiences affect risk taking?, *Quarterly Journal of Economics* 126, 373–416.

Malmendier, Ulrike, and Stefan Nagel, 2015, Learning from inflation experiences, *The Quarterly Journal of Economics* 131, 53–87.

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

Pool, Veronika K., Noah Stoffman, and Scott E. Yonker, 2012, No place like home: Familiarity in mutual fund manager portfolio choice, *The Review of Financial Studies* 25, 2563–2599.

Schroder, Mark, and Costis Skiadas, 1999, Optimal consumption and portfolio selection with Stochastic Differential Utility, *Journal of Economic Theory* 89, 68–126.

Seasholes, Mark S, and Ning Zhu, 2010, Individual investors and local bias, *The Journal of Finance* 65, 1987–2010.

Shepard, Roger N., 1987, Toward a universal law of generalization for psychological science, *Science* 237, 1317–1323.

Shive, Sophie, 2010, An epidemic model of investor behavior, *Journal of Financial and Quantitative Analysis* 45, 169–198.

Sims, Chris R., 2018, Efficient coding explains the universal law of generalization in human perception, *Science* 360, 652–656.

Stanton, Richard, Johan Walden, and Nancy Wallace, 2014, Securitization networks and endogenous financial norms in U.S. mortgage markets, SSRN Working Paper.

Stanton, Richard, Johan Walden, and Nancy Wallace, 2018, Mortgage loan flow networks and financial norms, *The Review of Financial Studies* 31, 3595–3642.

Trojani, Fabio, and Paolo Vanini, 2004, Robustness and ambiguity aversion in general equilibrium, *Review of Finance* 8, 279–324.

Vissing-Jorgensen, Annette, 2003, Perspectives on behavioral finance: Does "irrationality" disappear with wealth? Evidence from expectations and actions, *NBER Macroeconomics Annual* 139–193.

Walden, Johan, 2019, Trading, profits, and volatility in a dynamic information network model, *Review of Economic Studies* 86, 2248–2283.

Weil, Philippe, 1990, Nonexpected utility in macroeconomics, *Quarterly Journal of Economics* 105, 29–42.